CONVEX OPTIMIZATION METHODS FOR ROBUST AND DETERMINISTIC OPTIMAL POWER FLOW PROBLEMS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Raphael Louca

December 2017

CONVEX OPTIMIZATION METHODS FOR ROBUST AND DETERMINISTIC OPTIMAL POWER FLOW PROBLEMS

Raphael Louca, Ph.D.

Cornell University 2017

Today's energy sector faces a trilemma of challenges: keeping electricity reliable, affordable, and clean. Providing affordable and reliable service entails the determination of an operating point for the power system, which minimizes the total cost of generation while respecting the physical and operational constraints of the system. At the core of reliability and affordability challenges is the *AC optimal power flow problem (AC-OPF)*. In its most general form, the AC-OPF problem is a high-dimensional optimization problem that is nonconvex and NP-hard in general. In addition to reliability and affordability challenges, in recent years there has been a growing need to develop optimization methods, which enable the reliable and efficient operation of power systems that have a large fraction of their power supplied from intermittent renewable energy resources, like wind and solar. The need to accommodate the intrinsic uncertainty in the power supply of such resources will require the development of robust optimization methods for the AC-OPF problem. In its most general formulation, the *robust AC optimal power flow (RAC-OPF)* problem amounts to a two-stage robust optimization problem, in which the system operator must determine a day-ahead generation schedule that minimizes the expected cost of dispatch, given an opportunity for recourse to adjust its day-ahead schedule in real-time when the uncertain system variables have been realized. In addition to being nonconvex, the RAC-OPF problem is an infinite-dimensional optimization problem due to the need to optimize over an infinite-dimensional recourse policy space.

The central topic of this thesis is the development of computationally tractable convex inner and outer approximations (relaxations) for the AC-OPF problem and the RAC-OPF problem. In the first part of the thesis, we focus on the AC-OPF problem, its equivalent reformulation as a rank one constrained semidefinite program, and its semidefinite programming relaxation. First, we study an a priori sufficient condition developed in the literature, which guarantees the exactness of the relaxation and then we develop an posteriori sufficient condition, which can be used to verify the inexactness of the relaxation. For AC-OPF problems that do not satisfy the sufficient condition for exactness, we investigate the extent to which it is possible to apply a structured perturbation to the problem data to obtain a problem, which satisfies said sufficient condition and which yields an optimal solution that is feasible for the original problem. An explicit bound on the performance of the feasible solution is also derived. In addition to perturbation-based inner approximations, we also propose an inner approximation scheme, which is based on an equivalent representation of the rank one constraint as the difference of two convex functions. Using this representation, we develop an algorithm, which is guaranteed to generate a sequence of feasible solutions with nonincreasing costs. Lastly, we propose an iterative linearization-minimization algorithm to uncover rank one optimal solutions for the semidefinite relaxation when the relaxation is exact, but its optimal solution set contains both rank one and high rank optimal solutions. A simple bisection method is also proposed to address problems for which the linearization-minimization procedure fails to return a rank-one optimal solution.

In the second part of the thesis, we focus on the RAC-OPF problem. By restricting the space of recourse policies to those which are affine in the uncertain problem data, we propose a method to approximate RAC-OPF from within by a finite-dimensional semidefinite program. The solution of this optimization problem gives rise to an affine recourse policy for the RAC-OPF problem. In addition to the inner approximation, we develop a method for constructing a second-order cone outer approximation to the RAC-OPF problem. The crux of our approach centers on the reformulation of RAC-OPF as a robust rank one constrained semidefinite program, which is then relaxed to a robust linear program. The relaxation is obtained by eliminating the rank one constraint and by approximating the cone of positive semidefinite matrices from without by a polyhedral cone. A recursive method is also developed, which refines said polyhedral cone in those regions that are important for optimization to improve the performance of the relaxation. The practical value of our method is that one can obtain a feasible solution to the RAC-OPF problem by solving a finite-dimensional semidefinite program whose suboptimality can be bounded by solving a second-order cone program. In addition, if the gap between the optimal values of the outer and inner approximations is small, we have a certificate of near optimality of the feasible solution obtained. To the best of our knowledge, our approach is the first to provide a systematic method for computing a feasible solution to the RAC-OPF problem and a nontrivial global lower bound on its optimal value via convex optimization.

As the final contribution, we develop computationally tractable inner and outer approximations to robust semidefinite programs, a class of optimization problems that is NP-hard in general. The proposed method relies on approximating the positive semidefinite cone from within and without by appropriate polyhedral cones. In addition, we develop a recursive method, which refines these cones to sharpen the approximations. In particular our method, eliminates the optimal solution of the approximation at the current iteration step from the feasible set of the approximation at the next iteration step.

BIOGRAPHICAL SKETCH

Raphael Louca received a BSc degree in Mathematics and Electrical Engineering from Pennsylvania State University in University Park PA, in 2011. He completed his graduate work in Electrical and Computer Engineering at Cornell University in Ithaca NY, where he earned an MEng degree in May 2011, an M.S. degree in May 2016, and a PhD degree in December 2017.

In 2007, Raphael received a Fulbright Scholarship under the Cyprus America Scholarship Program (CASP) for undergraduate studies in the United States. He is also a recipient of the Jacobs and H.C. Torng fellowships from the School of Electrical and Computer Engineering at Cornell University and the Evi Sofianou fellowship from the Evi Sofianou Foundation.

His research interests lie in the intersection of power systems and optimization theory. His work has centered on characterizing efficient convex approximations for a problem in power systems referred to as optimal power flow. He has also worked on characterizing properties of solutions to sparse semidefinite programs.

To my parents, Klitos and Theodora Louca, my sister Maria Louca, and Annie.

For their unconditional love, support, and encouragement.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my PhD advisor, Professor Eilyan Bitar, for his guidance, patience, and mentorship throughout my graduate studies. His deep technical knowledge, originality, creativity, and elegance in scientific presentation have been inspirational and invaluable. Professor Bitar has instilled in me the importance of being adaptable and critical of my work and has encouraged me to never shy away from the details. He has impressed on me the value of precision both in writing and communicating. Beyond his mentorship, Professor Bitar has also invested in developing my skills by giving me the opportunity to attend several workshops related to my research interests. I owe him a lot and I will always carry with me his advice and guidance.

I would also like to thank Professor Peter Seiler with whom I had the opportunity to collaborate early in my PhD. I am thankful to him for sharing his invaluable knowledge in the theory of linear matrix inequalities and for pushing me to think outside the box. I have always looked up to Professor Seiler's mathematical rigor and logical thinking. My gratitude also goes to Professor Subhonmesh Bose, whom I met as a postdoctoral scholar at Cornell. I'll never forget our long walks and the plethora of ideas that emerged from them. Lastly, my deepest appreciation goes to Professors Adrian Lewis, James Renegar, and Michael Todd, for always providing sound advice and deep insight and for helping me conceptualize seemingly difficult mathematical programming concepts. I have learned more from taking convex analysis with Professor Lewis than from any other class in both my undergraduate and graduate studies.

I would also like to thank Ibrahim Issa, Alex Ivanov, Daniel Munoz Alvarez, Kia Khezeli, Weixuan Lin, and Yiting Xie who have been not only my collegues, but also my great friends. Talking things out with them – be it academic or personal – has been invaluable and they have always given me great and sound advice. I will never forget the "why is this true?" math questions Ibrahim and I asked each other at 3:00am in the morning, hours before the homework was due. Special thanks also go to my dear friend Barbara who has been very supportive and encouraging over these years. I will never forget the Sunday brunches with her and Ibrahim at the Statler. Barbara was my go-to person for everything Ithaca related and a constant reminder that taking time for myself can go a long way. In the five years at Cornell, I had the pleasure to meet and develop a friendship with many other people. Special thanks go to Kolbeinn, Katla, Levon, Tayyar, Deepan, Hazal, Rose, Nithin, Ilan, Alireza, Chiun Lin, Enrique, Sina, David, Andrew, Sergio, Angeligue, and Nirmal.

My family has been an inexhaustible well of encouragement and support throughout the five years of this journey. I couldn't thank my parents enough for supporting me, even though they were five thousand miles away physically. My parents instilled in me the importance of education from early on. More importantly, however, they instilled in me the importance of self-respect and respect for others and have constantly reminded me that these characteristics go hand-to-hand with education. Special thanks go to Maria (my sister) and Frixos for pulling me out of my head and for reminding me that respect is earned, not granted.

Lastly, I would like to wholeheartedly thank Annie, whom I met at a challenging stage in my life. Annie has been the counter voice to the critical, judgmental, and derogatory voice I had in my life at that stage. Annie was the person that helped me internalize my parent's urge for self-respect and she has been a constant reminder that there is more to a person than what meets the eye. Without her, this thesis would not have been possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Part I

Introduction

CHAPTER 1

TODAY'S POWER SYSTEMS: CHALLENGES AND DIRECTIONS

1.1 Reliable, Affordable, and Clean Power Systems

Electric power systems are an integral part of modern societies as they facilitate economic growth, promote business development, and improve our overall quality of life. In the United States, the electric power industry represents 3% of the real gross domestic product [33]. Today, power systems face a trilemma of challenges – keeping electricity reliable, affordable, and clean. Providing reliable service entails constant and prudent management of supply to meet the ever-changing system demand while satisfying both physical and operational constraints of the power network. Electricity is a commodity that we hardly ever think of, unless we do not have it. Service interruptions are extremely costly to society both economically and socially. For example, the August 2003 blackout left approximately fifty million people in the eastern United States in the dark for several days and had an estimated cost between four to ten billion dollars [53].

While keeping the lights on is of great importance, doing it in an efficient manner is challenging. As a matter of fact, supplying electricity in the least-cost manner while satisfying the power system's physical and operation constraints is a fundamental problem in power system operations, which is referred to as the *optimal power flow problem*. At its core, the optimal power flow problem, is an optimization problem that is computationally intractable, in general. The U.S. Energy Information Administration (EIA) estimates that electricity losses at both the transmission and distribution level account for about 5% of the electricity produced annually in the United States. In addition, according to the US

Federal Energy Planning Board (FERC), a mere 5% increase in the efficiency of algorithms for AC-OPF will yield six billion dollars in savings per year in the United States alone [18].

Until recently, the focus in power systems has been in keeping electricity reliable and affordable. In the last couple of decades, however, increased environmental concerns surrounding climate change have induced many U.S. states to adopt legislation mandating that a significant percentage of their electricity be generated by clean renewable resources. The state of California, for example, has set a target of 33% renewable energy penetration by the year 2020. One of the fundamental barriers to deep integration stems from the variability of power from wind and solar resources, which are highly intermittent, non-dispatchable, and difficult to forecast. As a consequence, they require costly reserve generation to firm their output. For example, the 2010 eastern wind integration and transmission study report by the National Renewable Energy Laboratory projects that reserve requirements will increase by 1500 MW under a 20% wind energy penetration scenario in the PJM interconnection [64]. If these increases are met with combustion-fired generation, they will be both economically untenable and counterproductive to carbon emissions reductions. As as result, the drive to support the deep integration of variable renewable energy into the grid, without sacrificing reliability, will require a paradigm shift in how we produce, deliver, and consume energy.

The drive to support reliable, affordable, and clean energy into the grid will require a paradigm shift in how we produce, deliver, and consume energy. To address these challenges, a wide body of research has recently emerged to create the so-called *smart grid*, which will enable increased control over the electric power grid by integrating advanced sensing and communications that will enhance the day-to-day operation of the grid. The smart grid is expected to improve the way utility companies manage physical assets, the way consumers interact with their energy supply, and the way governments reform policies in order to address the challenge of reducing greenhouse gas emissions.

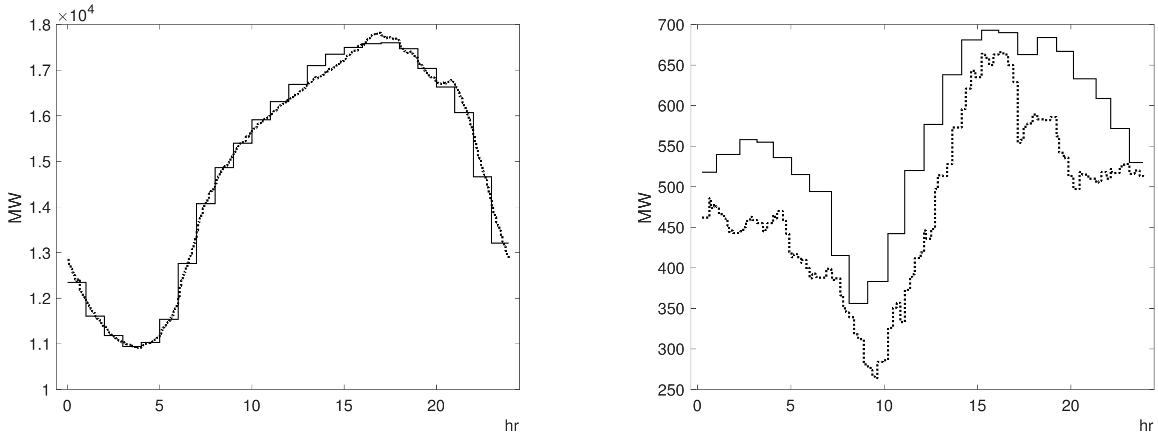
1.2 Sources of Uncertainty

1.2.1 Failures and Contingencies

In a system consisting of hundreds of thousands of components, the failure of one component is not a rare event. Such events are exacerbated by the aging system infrastructure and by the exposure of certain components (such as transmission lines) to inclement weather conditions. The social cost of power outages can be large, and it is agreed that power systems must be able to withstand plausible disturbances and operate at new states long enough to give the system operator enough time to restore the system to its normal operating point.

1.2.2 Load

When aggregated over a large number of consumers, demand forecasts are usually more accurate than individual disaggregate forecasts, as they tend to have smaller standard deviation of error relative to the mean. In addition, such aggregate demand forecasts are naturally correlated with weather conditions – like ambient temperature and humidity – which affect the heating and cooling loads of buildings. One-to-four day weather forecasts are typically accurate within a few percent. As a result, independent system operators



(a) (b)

Figure 1.1: Figures (a) and (b) depict the day-ahead and real-time (solid line) forecasts for load and wind, respectively, in the ISO New England for the operating day of July 6th, 2017.

(ISOs) can rely on such forecasts to produce fairly accurate aggregate demand forecasts (see Figure 1.1(a)). In addition to weather conditions, demand electricity consumption exhibits strong periodical patterns. For example, electricity consumption is at peak in the morning hours when consumers arrive at work and in the evening when consumers return home. In addition, electricity consumption during weekends and holidays tends to be lower.

1.2.3 Supply

Today's electric power systems rely on conventional generators using fossil fuels (e.g., coal, oil, natural gas) to produce power in order to balance the system's demand. Such generation resources are dispatchable, giving the system operators the flexibility to dynamically tailor the dispatch schedule of generators to compensate unforeseen fluctuations in demand. Therefore, the system is said to be operated under a *supply-follows-demand* paradigm. Recently, the drive to support the deep integration of variable renewable energy resources into the grid, without sacrificing reliability, has brought unprecedented challenges to power system operations. Renewable energy resources, like wind and solar, are non-dispatchable, highly intermittent, and difficult to forecast. This is a fundamental barrier to their large-scale integration and currently system operators must maintain costly reserve generation to support the grid when renewable generation levels differ significantly from their forecasted levels (see Figure 1.1(b)). As a result, with the increased penetration of renewable energy resources, we will witness a loss of supply-side flexibility.

In contrast to demand, the maximum available capacity of intermittent renewable resources is difficult to forecast within a few percent accuracy, even on relatively short timescales. In the case of wind and solar, for example, this difficulty derives from the nonlinear relationship between the instantaneous power produced from such resources and several weather attributes – like wind speed and incident radiation – which can be foretasted fairly accurately. In particular, the instantaneous power produced by solar panels depends on the panel's orientation and its angle of tilt with the horizontal. Both of these parameters affect the solar radiation reaching the panel's surface and, in turn, the power produced by the panel.

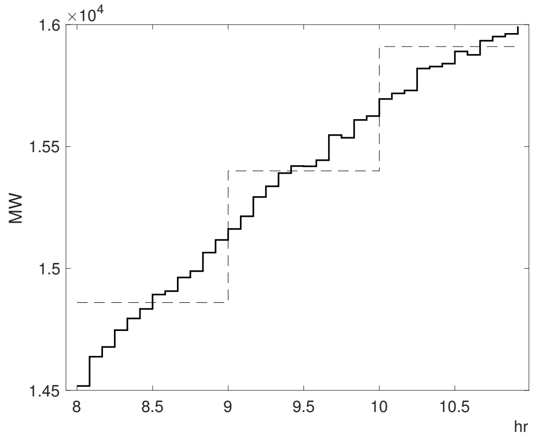
Today, wind and solar power production are assimilated into the grid through legislative mandates. In California, in particular, the ISO must accept all produced wind power and therefore, wind power is treated as a negative load. This yields an increase in the variability of net load (the total demand in the system minus renewable energy generation) which is absorbed by reserve generators. The cost associated with deploying such generators is beared by the load serving entities.

1.3 Two Settlement System

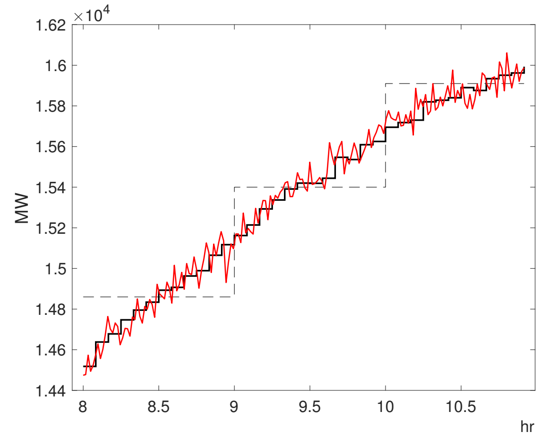
Wholesale electricity markets predominantly run on a two settlement market system, which consists of a day-ahead (DA) forward market and a real-time (RT) spot market. In day-ahead, the system operator must schedule an initial dispatch of its resources subject to uncertainty in the eventual realization of certain system variables, including demand and generation levels of renewable resources. Such DA scheduling decisions are essential, as certain generation resources (e.g., coal and nuclear) have limited ramping capabilities. In real-time, all uncertain variables are realized, and the system operator is provided a recourse opportunity to adjust its DA dispatch schedule in order to balance the system at minimum cost.

1.3.1 Day-Ahead Dispatch

The day-ahead (DA) market takes place the day before the operating day, and it is aptly named. In the DA market, suppliers submit production bids for delivery of power the following day. These bids take the form of price/quantity pairs for delivery of constant power over some time interval, typically, of length one hour (see Figure 1.2(a)). After collecting all the bids, the system operator schedules an initial dispatch of the generators to meet the forecasted demand. This is referred to as *market clearing*. Producers, whose bids have been accepted, are bound to a financial obligation and are penalized for deviations from their day-ahead schedules. The day ahead market enables the dispatch of generators in real-time which have low marginal costs, but long start-up or ramp times. Therefore, the more efficient the scheduling of such generating resources, the lower the cost of power.



(a)



(b)

Figure 1.2: Figure 1.2a depicts the day-ahead (dashed line) hourly dispatch schedule and real-time (solid line) five-minute dispatch schedule for the supply of power in the ISO New England for the operating day of July 6, 2017. Figure 1.2b depicts the realized load. Any residual imbalance that occurs between the real-time dispatch schedule and the realized load is compensated by automatic control mechanisms.

We emphasize that the day-ahead market is a financial, not a physical market. It is in the real-time market that generators produce power to keep the lights on.

1.3.2 Real-time Dispatch

The day-ahead market is cleared well in advance of the operating day, therefore, in order to ensure that supply and demand are balanced, a real-time market is employed. In the real-time market, suppliers can adjust their day-ahead schedules based on new information such as updated renewable energy, price, and load forecasts. Contrary to the day-ahead market, the real-time market is a physical market for delivery of constant power over fifteen or five minute time intervals (see Figure 1.2(a)). Any residual imbalance that occurs between supply and demand at smaller time scales (sub-seconds to minutes) is compensated by automatic control mechanisms. These control mechanisms are responsible for fast adjustments to the output of generators, which are necessary for keeping supply and demand balanced at all times and for maintaining certain system parameters (e.g. frequency and voltage profile) within design limits.

1.4 Unit Commitment

Certain generation resources (or units) have limited temporal flexibilities, e.g., long start-up and shut-down times and minimum online and offline times. Utilities make "commitment" decisions day to hours in advance to have certain generation resources available to produce power, when the need arises. Such commitment decisions mean that utilities are willing to incur fixed costs associated with generators' start-up costs in order to have these resources available to produce power in real time. These decisions are financially justified since these generation resources tend to have lower marginal costs than generation resources with faster start-up times (e.g., gas plants).

The objective of the *unit commitment problem* is to find a dispatch schedule which minimizes both the commitment and dispatch costs of meeting demand. The decision variables in the unit commitment problem are the statuses of the generator units, which are inherently discrete, and their output levels which are inherently continuous. The resulting optimization problem is a mixed-integer programming problem.[1] Due to its nonconvexity and large scale, the unit commitment problem poses computational challenges. We emphasize that this thesis does not consider the unit commitment problem. Namely, we

[1]A mixed integer programming problem is an optimization problem, which includes both descrete (integer) and continuous (real or complex) decision variables.

assume throughout that the on and off statuses of the generators are known.

1.5 Optimal Power Flow

A closely related problem to the unit commitment problem is the *optimal power flow problem*. This problem is commonly referred to as the economic dispatch problem. In contrast to unit commitment, in the optimal power flow problem, the on and off statuses of the generators are taken to be known. The objective is to find a dispatch for the collection of generators that are committed to minimize the cost of meeting demand, while respecting physical and operational constraints. The physical constraints represent the power balance equations described by Kirchhoffs current and voltage laws, while the operational constraints reflect bounds on real and reactive power generation, branch flows, and voltage magnitudes. In contrast to unit committment, optimal power flow involves only continuous decision variables, i.e., generation output and voltage phasors and it is often formulated as a single period optimization problem. In any given day, the optimal power flow problem is solved every five to fifteen minutes (real-time dispatch) to account for fluctuating demands and changes in operating conditions.

1.5.1 Challenges

In its most general formulation, the optimal power flow problem is a large scale, nonconvex optimization problem that is NP-hard. The high dimensionality of the optimal power flow problem derives from the existence of a decision variable representing the voltage phasor at each bus in the power system. The nonconvexity arises because the feasible set has a nonconvex quadratic dependency on the set of complex bus voltages. Because of this non-convexity, the optimal power flow problem may admit several locally optimal solutions – some of which may be suboptimal. Since its origin, a variety of techniques from mathematical programming, including linear and quadratic programming, have been proposed to solve the optimal power flow problem. For a comprehensive literature survey, the interested reader is referred to [18] and to the references therein. In practice, the predominant approach to solving the optimal power flow problem involves the implementation of nonlinear optimization routines, capable of addressing the inherent non-convexity (e.g. MATPOWER [3], PSSE). These solvers, however, do not offer any guarantees regarding the global optimality of the solution they produce. According to the US Federal Energy Planning Board (FERC), it is estimated that a mere 5% increase in the efficiency of algorithms for optimal power flow will yield six billion dollars in savings in the United States alone [18].

The drive to support the deep integration of renewable energy resources into the grid, without sacrificing reliability has brought additional challenges to the optimal power flow problem. In particular, the need to accommodate the intrinsic uncertainty in the power supply of such resources will require the development of robust optimization methods for the optimal power flow problem. In addition to being nonconvex and high dimensional, the robust optimal power flow problem is an infinite-dimensional optimization problem due to the need to optimize over an infinite-dimensional recourse policy space and to enforce an infinite number of constraints.

1.6 Summary of Contributions and Dissertation Organization

In this thesis, we study the effectiveness of convex approximations to the alternating current optimal power flow problem (AC-OPF). The results presented in the thesis are organized according to two main parts: one in which all system parameters are known and one in which certain system parameters, such as the maximum available capacities of generation resources, are uncertain. In the first part, the AC-OPF problem is formulated as a single-stage static optimization problem and the objective of the ISO is to determine an operating point for the power system, which minimizes the total cost of generation. In the second setting, the AC-OPF problem is formullated as a two-stage robust optimization problem (henceforth, we refer to this problem as RAC-OPF), in which the system operator must determine a day-ahead generation schedule, which minimizes the expected cost of dispatch, given a recourse opportunity to adjust its day-ahead schedule in real-time when the uncertain system variables are realized, e.g., the available supply from renewable resources. The RAC-OPF problem allows the system operator to compensate the attending intermittency and uncertainty in the supply of power brought by renewable energy resources to the power system operation. The dissertation is organized as follows.

In Chapter 2, Alternating Current (AC) Optimal Power Flow, we formulate the classical AC-OPF problem, which takes the form of a complex quadratically constrained quadratic program, and study its second-order cone and semidefinite relaxations. We discuss an a priori sufficient condition, which guarantees exactness of the semidefinite relaxation and develop an a posteriori sufficient condition to verify inexactness of the semidefinite relaxation. In addition, we discuss nonconvex optimization approaches based on primal-dual interior point methods and the alternating direction method of multipliers for obtaining locally optimal solutions to the AC-OPF problem. The first is based on a primal-dual interior point method and the second on the alternating direction method of multipliers.

In Chapter 3, Perturbation Methods for Quadratically Constrained Quadratic Programs, we consider complex quadratically constrained quadratic programs (QCQPs) having arbitrary graph structures. Recent results have shown that QCQPs having acyclic graph structures can be solved in polynomial time, provided that their constraints satisfy a certain technical condition. We investigate the extent to which it is possible to apply structured perturbations on the problem data to yield acyclic QCQPs having optimal solutions satisfying certain approximation guarantees. Specifically, we provide sufficient conditions under which the perturbed QCQP can be solved in polynomial time to yield a feasible solution to the original QCQP and derive an explicit bound on the performance of said solution in the worst case.

In Chapter 4, Recursive Semidefinite Approximations of AC-OPF, we consider the reformulation of the AC-OPF problem as a semidefinite program with a rank one constraint on the set of feasible matrices. We provide an equivalent representation of the rank one inequality as the difference of two convex functions. Using this representation, we develop a convex inner approximation to the AC-OPF problem as a semidefinite program and propose an algorithm, which yields a sequence of feasible solutions with nonincreasing costs. In addition to the convex inner approximation, we consider the semidefinite relaxation of the AC-OPF problem and propose an iterative linearization-minimization algorithm to uncover hidden rank one optimal solutions to the relaxation in case its optimal solution set contains both both high rank and rank one matrices. A simple bisection method is also proposed to address problems for which the linearization-minimization procedure fails to yield a rank-one optimal solution.

In Chapter 5, Robust AC Optimal Power Flow, we formulate the robust AC optimal power flow (RAC-OPF) problem as a two-stage robust optimization problem with recourse. This problem amounts to an infinite-dimensional nonconvex optimization problem, which is computationally intractable, in general. By restricting the space of recourse policies to those which are affine in the uncertain problem data, we provide a method to approximate RAC-OPF from within by a finite-dimensional semidefinite program. The resulting semidefinite program yields an affine recourse policy that is guaranteed to be feasible for RAC-OPF. In addition to the inner approximation, we develop a method for constructing a conic outer approximation to the RAC-OPF problem. The crux of our approach centers on the reformulation of the RAC-OPF problem as a robust rank one constrained semidefinite program and on its relaxation to a robust linear program. This relaxation is obtained by eliminating the rank constraint and by approximating the cone of positive semidefinite matrices from *without* by a polyhedral cone. We also develop a recursive method, which refines the polyhedral cone only on those regions that are important for optimization in order to improve the performance of the relaxation. The practical value of our approximation techniques proposed in this Chapter derive from the fact that one can obtain a feasible solution to the RAC-OPF problem by solving a finite-dimensional semidefinite program; and can bound the suboptimality incurred by this feasible solution by solving another finite-dimensional conic linear program. And if the gap between the optimal values of the outer and inner approximations is small, we have an a posteriori certificate of near optimality of the feasible solution obtained.

Finally, in Chapter 6, Recursive Conic Approximations of Robust Semidefinite Programs, we investigate the problem of approximating solutions to intractable robust semidefinite programs. The proposed method relies on approximating the positive semidefinite cone from within and without by appropriate polyhedral cones. In addition, we develop a recursive method, which adaptively refines these cones to sharpen the approximations. The proposed method is shown to eliminate the optimal solution of the approximation at the current iteration step from the feasible set of the approximation at the next iteration step.

Part II

Deterministic AC Optimal Power Flow

CHAPTER 2

AC OPTIMAL POWER FLOW

2.1 Introduction

The alternating current optimal power flow (AC-OPF) problem is a classic problem in power systems operations that has been studied extensively beginning with the seminal work of Carpentier [17] in 1962. The AC-OPF problem is generally formulated as a static optimization problem where the objective is to minimize a convex cost function subject to possibly nonconvex physical and operational constraints. The cost function is typically chosen to represent either the total cost of generation, line power losses, or the sum of voltage magnitudes across transmission buses. The cost is assumed to be affine or convex quadratic. The physical constraints represent the power balance equations described by Kirchhoff's current and voltage laws, while the operational constraints reflect bounds on real and reactive power generation, branch flows, and voltage magnitudes. Commonly, the set of decision variables are comprised of a combination of bus complex power injections and voltages. Naturally, the solution to AC-OPF is given by a set of decision variables that yield a minimal cost operating point of the power system. Although AC-OPF is straightforward to formulate, it is in general difficult to solve.

In its most general formulation, the AC-OPF problem is a high dimensional, nonconvex optimization problem that is NP-hard. The nonconvexity arises because the feasible set has a nonconvex quadratic dependency on the set of complex bus voltages. Because of this nonconvexity, the AC-OPF problem may admit several locally optimal solutions – some of which may be suboptimal. Since its origin, a variety of techniques from mathematical programming, including linear and sequential quadratic programming, have been proposed to solve the AC-OPF problem. We review some of these approaches in Sections 2.4-2.6. For a comprehensive literature survey, the interested reader is referred to [18] and to the references therein. In practice, the predominant approach to solving AC-OPF involves the implementation of nonlinear optimization routines, capable of addressing the inherent nonconvexity in the problem (e.g., MATPOWER [96]). These solvers, however, do not offer any guarantees regarding the global optimality of the solution they produce.

More recently, there has been a flurry of work exploring the use of convex relaxations for solving the AC-OPF problem [4, 34, 41]. In particular, the second-order cone and the semidefinite relaxations have garnered considerable attention. Qualitatively, these relaxations involve first recasting AC-OPF as nonconvex quadratically constrained quadratic program (QCQP) and then reformulating the nonconvex QCQP as a semidefinite program with a rank one inequality constraint on the set of feasible matrices. Both relaxations entail removing the rank one constraint to obtain a semidefinite program. In addition to the rank constraint, the second-order cone approach further relaxes the positive semidefiniteness constraint. Namely, it requires only the two-by-two principal minors of all feasible matrices to be nonnegative.[1] For networks with acyclic topologies, both relaxations are shown to be equivalent [14]. The relaxations are said to be *exact* if their optimal solution set contains a rank one positive semidefinite matrix – a condition which is difficult to verify in practice. In this chapter, we focus primarily on the semidefinite relaxation of the AC-OPF problem.

Certain realizations of AC-OPF yield semidefinite relaxations, which have optimal solutions of rank no greater than one [41, 94, 13]. However, it has been observed that in practice the

[1]A matrix is positive semidefinite if and only if all its principal minors are nonnegative

semidefinite relaxation of some AC-OPF problems have optimal solutions of high rank – even though rank one optimal solutions may exist. This raises several interesting questions. For instance, when is the minimal rank of the optimal solution set of the semidefinite relaxation strictly greater than one? Alternatively, in situations where the optimal solution set contains matrices of multiple rank, how might one uncover a *hidden* rank one optimal solution when it exists? In Section 2.5.1 we develop an a posteriori sufficient condition, which addresses the former question. And in Section 2.5.2 we provide a priori sufficient condition proposed in [13, 76], which addresses the latter question.

Related work: After reformulating the complex QCQP describing the AC-OPF problem as a QCQP over real-valued decision variables, Lavaei and Low [41] propose solving the dual relaxation of said problem and provide a sufficient condition under which the solution to the relaxed problem will be *globally optimal* for the original nonconvex problem. Their main theoretical result states that the duality gap is zero for the QCQP over the real-valued decision variables if the dual multiplier corresponding to the positive semidefinitness constraint has a zero eigenvalue of multiplicity two. The authors empirically observe that this condition is satisfied by many IEEE benchmark networks. However, several examples were given in [43] that demonstrate the failure of semidefinite relaxations to yield rank one optimal solutions in the case of networks with binding line flow constraints realizing negative locational marginal prices.

Building on this work, Zhang and Tse [94] explore as to whether the relaxation is exact for certain families of networks. The authors show that for tree topologies satisfying certain constraints on the nodal power injections, the set of feasible active power injections and its convex hull have the same Pareto frontier. Therefore, the minimization of an increasing function over the convex hull of the feasible set will yield solutions on the Pareto frontier of the non-convex problem. Moreover for linear objectives, they claim that the semidefinite relaxation will yield a unique rank one optimal solution. Bose et al., build on these results by showing that nonconvex QCQPs having an underlying acyclic structure and satisfying certain technical conditions will yield semidefinite relaxations obtaining rank one optimal solutions (cf. [13] and Section 2.5.1). For general problem structures, however, the semidefinite relaxation may fail to yield optimal solutions that can be efficiently mapped back to the original feasible set.

Contribution: The primary contribution of this Chapter is the development of an posteriori sufficient condition for the nonexistence of rank-one optimal solutions to semidefinite relaxations of AC-OPF. This sufficient condition, which is presented in Section 2.5.2, exploits dual nondegeneracy of semidefinite programs as defined by Alizadeh et al. in [1] to establish uniqueness of primal optimal solutions to complex semidefinite programs. Of import is the fact that this sufficient condition holds for arbitrary network topologies, including trees.

Organization: The remainder of the Chapter is organized as follows. In Section 2.2 we formulate the classical AC-OPF problem and then in Section 2.3 we reformulate AC-OPF as a complex quadratically constrained quadratic program. In Section 2.4, we develop the second-order cone and semidefinite relaxations of AC-OPF. Section 2.5.1 presents a sufficient conditions developed in the literature, which guarantees the exactness of the semidefinite relaxation. In Section 2.5.2, we develop an posteriori sufficient condition, which can be used to verify that the semidefinite relaxation is inexact. Section 2.6 presents two nonconvex optimization approaches developed in the literature for obtaining locally optimal solutions to the AC-OPF problem. The first is based on a primal-dual interior point method and the second on the alternating direction method of multipliers. Conclusions are given in Section 2.7.

Notation: Let $\mathbf{R}$ be the field of real numbers and $\mathbf{C}$ the field of complex numbers. For $z \in \mathbf{C}$, let $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$ be the real and imaginary parts of z, respectively. In addition, let $\mathbf{i} := \sqrt{-1}$ be the imaginary unit. Let $\mathbf{R}^n$ be the n-dimensional real vector space and $\mathbf{H}^n$ be the space of $n \times n$ Hermitian matrices. Given a matrix A, let $[A]_{ij}$ be its (i,j) entry. And denote by $A^\top$ and A^* the transpose and complex conjugate transpose of A, respectively. For a matrix $A \in \mathbf{H}^n$, the notation $A \succeq 0$ $(A \succ 0)$ means that A is positive semidefinite (positive definite). Endow $\mathbf{R}^n$ with the inner product $x^\top y$ for all $x, y \in \mathbf{R}^n$ and $\mathbf{H}^n$ with the trace inner product $\mathrm{tr}(X^* Y)$ for all $X, Y \in \mathbf{H}^n$.

2.2 Formulation of Deterministic AC-OPF

We begin with a development of a general model for the AC optimal power flow problem (AC-OPF). The perspective we adopt is that of the system operator, whose objective is to determine the dispatch of generation resources in order to minimize the cost of meeting demand, while ensuring that all operational limits of generation and transmission facilities are met.

We consider an electric power network whose topology is described by an undirected graph $\mathcal{G} := (\mathcal{V}, \mathcal{E})$, where the vertex set $\mathcal{V} := \{1, \ldots, n\}$ represents the collection of transmission buses and the edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ represents the collection of transmission lines connecting buses. We refer the reader to [25] for background on graph theory. We describe the AC power balance equations, which govern the relationship between complex bus voltages and power injections [9] (we refer the reader to Appendix A, where we provide a detailed derivation of these equations). Let $Y \in \mathbf{C}^{n\times n}$ be the network admittance matrix, $v \in \mathbf{C}^n$ the vector of bus voltages, and $s \in \mathbf{C}^n$ the vector of complex bus power injections. The AC power balance equations can be expressed as

$$s_i := v^* S_i v, \qquad i \in \mathcal{V}, \tag{2.1}$$

where $s_i \in \mathbf{C}$ denotes the net complex power injected at bus i and $S_i := Y^* e_i e_i^* \in \mathbf{C}^{n\times n}$. The complex power flow from bus i to bus j is denoted by $s_{ij} \in \mathbf{C}$ and is given by

$$s_{ij} := v^* S_{ij} v, \qquad (i,j) \in \mathcal{E}, \tag{2.2}$$

where $S_{ij} := e_i e_i^* (\widehat{y}_{ij}/2 - [Y]_{ij})^* + e_j e_i^* [Y]_{ij}^* \in \mathbf{C}^{n\times n}$ and $\widehat{y}_{ij} \in \mathbf{C}$ denotes the total shunt admittance of line $(i,j) \in \mathcal{E}$. We enforce two classes of constraints. The first requires that bus voltage magnitudes satisfy

$$v_i^{\min} \leq |v_i| \leq \ v_i^{\max}, \qquad i \in \mathcal{V},$$

where $v_i^{\min},\ v_i^{\max} \in \mathbf{R}$ denote upper and lower bounds on the voltage magnitude at bus $i \in \mathcal{V}$. The second class of constraints enforce line flow capacities. Namely, the real power flow from bus i to bus j must satisfy

$$-\ell_{ij}^{\max} \leq v^* P_{ij} v \leq \ell_{ij}^{\max}, \qquad (i,j) \in \mathcal{E}, \tag{2.3}$$

where $P_{ij} := (S_{ij} + S_{ij}^*)/2 \in \mathbf{H}^n$ and $\ell_{ij}^{\max} \in \mathbf{R}$ denotes the real power flow capacity of line (i,j).

Let $d_i \in \mathbf{C}$ denote the power demand at bus i. If bus i is not connected to a load, we set $d_i = 0$. Denote by $\mathcal{V}_g \subseteq \mathcal{V}$, the subset of buses connected to generators. The power balance equations at load buses (i.e., buses not connected to generators) satisfy

$$s_i = -d_i, \quad i \in \mathcal{V} \setminus \mathcal{V}_g.$$

Similarly, at generator buses, the power balance equations satisfy

$$s_i = g_i - d_i, \quad i \in \mathcal{V}_g. \tag{2.4}$$

The power produced at each bus $i \in \mathcal{V}_g$ is constrained by the power capacity of each generator, namely,

$$g_i^{\min} \le g_i \le g_i^{\max}, \quad i \in \mathcal{V}_g,$$

where $g_i^{\min} \in \mathbf{C}$ and $g_i^{\max} \in \mathbf{C}$ denote the minimum and maximum nameplate power capacities of generator $i \in \mathcal{V}_g$. Moreover, each generator i, incurs a cost, which is assumed to be linear in the real power produced. We explicitly define its production cost as

$$f_i(g_i) \coloneqq \alpha_i \mathrm{Re}\{g_i\}, \tag{2.5}$$

where $\alpha_i \ge 0$ for all $i \in \mathcal{V}_g$. To maintain clarity of exposition, we assume that there is at most a single load and at most a single generator at each bus $i \in \mathcal{V}$.

Leveraging on the preceding development, we formulate the AC-OPF problem as follows:

$$\begin{array}{lll} \underset{g_i \in \mathbf{C},\ v \in \mathbf{C}^n}{\text{minimize}} & \sum_{i \in \mathcal{V}_g} f_i(g_i) & \\ \text{subject to} & \underline{g}_i \le g_i \le \overline{g}_i & i \in \mathcal{V}_g, \\ & v^* S_i v - g_i = -d_i, & i \in \mathcal{V}_g, \\ & v^* S_i v = -d_i, & i \in \mathcal{V} \setminus \mathcal{V}_g, \\ & v_i^{\min} \le |v_i| \le v_i^{\max}, & i \in \mathcal{V}, \\ & |v^* P_{ij} v| \le \ell_{ij}^{\max}, & (i,j) \in \mathcal{E}. \end{array} \tag{2.6}$$

2.3 Quadratically Constrained Quadratic Programming Formulation

In this section, we reformulate the AC-OPF problem (2.6) as a quadratically quadratically constrained quadratic program. We do so by first eliminating the generator dispatch variables $g_i \in \mathbf{C},\ i \in \mathcal{V}_g$ through their direct substitution according to the nodal power balance equations (2.4). These transformations give rise to the following reformulation of (2.6).

$$
\begin{aligned}
\underset{v \in \mathbf{C}^n}{\text{minimize}} \quad & v^* \left(\sum_{i \in \mathcal{V}_g} \alpha_i \Phi_i \right) v + \sum_{i \in \mathcal{V}_g} \alpha_i \mathrm{Re}\{d_i\} \\
\text{subject to} \quad & \mathrm{Re}\{\underline{g}_i - d_i\} \le v^* \Phi_i v \le \mathrm{Re}\{\overline{g}_i - d_i\}, && i \in \mathcal{V}_g, \\
& \mathrm{Im}\{\underline{g}_i - d_i\} \le v^* \Psi_i v \le \mathrm{Im}\{\overline{g}_i - d_i\}, && i \in \mathcal{V}_g, \\
& v^* \Phi_i v = -\mathrm{Re}\{d_i\}, && i \in \mathcal{V} \setminus \mathcal{V}_g \\
& v^* \Psi_i v = -\mathrm{Im}\{d_i\}, && i \in \mathcal{V} \setminus \mathcal{V}_g \\
& (v_i^{\min})^2 \le v^* (e_i e_i^\top) v \le (v_i^{\max})^2, && i \in \mathcal{V}, \\
& -\ell_{ij}^{\max} \le v^* P_{ij} v \le \ell_{ij}^{\max}, && (i,j) \in \mathcal{E},
\end{aligned}
\tag{2.7}
$$

where the Hermitian matrices $\Phi_i, \Psi_i \in \mathbf{H}^n,\ i \in \mathcal{V}$ are given by

$$
\Phi_i := \frac{S_i + S_i^*}{2}, \qquad \Psi_i := \frac{S_i - S_i^*}{\mathbf{i}2}. \tag{2.8}
$$

We remark that $v^\star \Phi_i v = \mathrm{Re}\{s_i\}$ and $v^* \Psi_i v = \mathrm{Im}\{s_i\}$ denote active and reactive net power injections at bus i, respectively.

Remark 1 (Nonconvexity)**.** *The nonconvexity in AC-OPF* (2.7) *arises because the feasible set has a nonconvex quadratic dependency on the set of complex bus voltages. This follows since the matrices $\Phi_i, \Psi_i,\ i \in \mathcal{V}$ and $P_{ij},\ (i,j) \in \mathcal{E}$ are indefinite, in general.*

Problem (2.7) can be posed as an instance of the following general class of nonconvex quadratically constrained quadratic programs.

$$\begin{aligned} \underset{v \in \mathbf{C}^n}{\text{minimize}} \quad & v^* A_0 v \\ \text{subject to} \quad & v^* A_i v \leq b_i, \qquad i = 1, \ldots, m \\ & v^* A_i v = b_i, \qquad i = m+1, \ldots, m+\ell, \end{aligned} \tag{2.9}$$

where $A_0, A_1, \ldots, A_{m+\ell} \in \mathbf{H}^n$ and $b_1, \ldots, b_{m+\ell} \in \mathbf{R}$. For the AC-OPF problem (2.7), the number of inequality constraints is given by $m := 2(|\mathcal{V}| + |\mathcal{E}| + |\mathcal{V}_g|)$ and the number of equality constraints is given by $\ell := 2|\mathcal{V} \setminus \mathcal{V}_g|$.

2.4 Convex Relaxations of AC-OPF

There is rich theory which considers convex approximations to the AC-OPF problem, the most popular of which is the Direct Current Optimal Power Flow approach (DC-OPF). This approach relies on a linearization of the quadratic power flow equations which are justified by several physical properties of power flows in typical power systems, e.g., the voltage magnitudes at each bus are approximately equal to one per unit. These approximations of the power flow equation give rise to a linear programming approximation of the AC-OPF problem. While solutions to DC-OPF are useful in many instances (e.g. as initial conditions to several optimization algorithms), they may not be feasible for the original AC-OPF problem. Moreover, the optimal value of the DC-OPF approximation is neither an upper or a lower bound to the optimal value of the original AC-OPF problem. In contrast, convex relaxations rely on enlargements of the solution set of AC-OPF and therefore they offer the ability to check whether a solution is feasible (in fact globally optimal) for the original problem. If not feasible, then the optimal value to the convex relaxation provides a lower bound to the optimal value of the AC-OPF problem. And if the feasible set of a convex relaxation is empty, then this is a certificate that the original AC-OPF problem is infeasible.

In this section, we focus on two convex relaxations to the AC-OPF problem: a second-order cone and a semidefinite relaxation. We refer the reader to Appendix A.2.1 for a detailed derivation of the DC power flow model.

2.4.1 Semidefinite Programming Relaxation

A recent stream of work has explored the application of semidefinite relaxations to solve the AC-OPF problem (2.6). Essentially, this convex relaxation entails the reformulation of the nonconvex quadratically constrained quadratic program (2.9) as a rank one constrained linear program over the cone of Hermitian positive semidefinite matrices. The semidefinite relaxation of AC-OPF is obtained by removing the rank one constraint. In this section, we provide a detailed exposition on this relaxation approach. We also refer the reader to [48, 49] for a comprehensive survey of these results.

The reformulation of the QCQP (2.9) as a rank one constrained semidefinite program is obtain through a lifting procedure. In particular, for any matrix $B \in \mathbf{H}^n$, the scalar

$$v^* B v = \text{tr}(v^* B v) = \text{tr}(B v v^*),$$

where the last equality follows by the *cyclic property* of the trace operator.[2] Letting $V := vv^*$ be a positive semidefinite rank one matrix, we obtain a reformulation of AC-OPF (2.9) as a rank one constrained semidefinite program. The semidefinite relaxation entails removing

[2]The *cyclic property* of the trace states that $\text{tr}(BC) = \text{tr}(CB)$, for any two matrices $B, C \in \mathbf{H}^n$.

the rank one constraint and it is given by

$$\begin{array}{lll} \underset{V \in \mathbf{H}^n}{\text{minimize}} & \operatorname{tr}(A_0 V) & \\ \text{subject to} & \operatorname{tr}(A_i V) \leq b_i, & i = 1, \ldots, m, \\ & \operatorname{tr}(A_i V) = b_i, & i = m+1, \ldots, m+\ell, \\ & V \succeq 0. & \end{array} \tag{2.10}$$

The optimal value to the semidefinite relaxation serves as a lower bound the optimal value of AC-OPF. Moreover, if the optimal solution of (2.10) has rank one, then an optimal solution to AC-OPF can be constructed through an eigenvalue decomposition of said optimal solution. In particular, if $V^\star$ is a rank one optimal solution to (2.10) and $V^\star = U \Lambda U^*$ is an eigenvalue decomposition of $V^\star$, then

$$v = \sqrt{[\Lambda]_{11}} U e_1$$

is an optimal solution to the AC-OPF problem (2.9).

The relaxation is said to be *exact* if its optimal solution set contains a rank one matrix – a condition which is difficult to verify in practice. Certain realizations of AC-OPF yield semidefinite relaxations with optimal solutions of rank no greater than one. It has, however, been observed that in practice many instances of AC-OPF yield semidefinite relaxations with optimal solutions of high rank – even though rank one optimal solutions may exist. This raises several interesting questions. For instance, when is the minimal rank of the optimal solution set of the semidefinite relaxation strictly greater than one? We address this question in Section 2.5.2. Alternatively, in situations where the optimal solution set contains matrices of multiple rank, how might one uncover a hidden rank one optimal solution when it exists? We address this question in Section 2.5.1 and Chapter 4.

2.4.2 Second-Order Cone Programming Relaxation

The feasible set of the semidefinite relaxation of AC-OPF can be further relaxed by replacing the positive semidefiniteness constraint with a less stringent constraint. Namely, we require only that the two-by-two principal minors of all feasible matrices to be nonnegative. This restriction gives rise to a second-order cone programming relaxation for AC-OPF. The reduction to a second-order cone program follows since any Hermitain two-by-two matrix X is positive semidefinite if and only if

$$[X]_{11}, [X]_{22} \geq 0 \quad \text{and} \quad \big|[X]_{12}\big|^2 \leq [X]_{11}[X]_{22}.$$

And the above constraints can be equivalently expressed as second-order cone constraints of the following form:

$$[X]_{11}, [X]_{22} \geq 0 \quad \text{and} \quad \begin{bmatrix} [X]_{11} + [X]_{22} \\ 2[X]_{12} \\ [X]_{11} - [X]_{22} \end{bmatrix} \in \mathbf{L}^3,$$

where $\mathbf{L}^k := \{(t, x) \in \mathbf{R} \times \mathbf{C}^{k-1} \mid \|x\|_2 \leq t\}$ is the second-order (Lorentz) cone of dimension k. Leveraging on the preceding development, the second-order cone relaxation of the AC optimal power flow problem is given by

$$\begin{array}{lll} \underset{V \in \mathbf{H}^n}{\text{minimize}} & \text{tr}(A_0 V) & \\ \text{subject to} & \text{tr}(A_i V) \leq b_i, & i = 1, \ldots, m, \\ & \text{tr}(A_i V) = b_i, & i = m+1, \ldots, m+\ell, \\ & \text{tr}(e_i e_i^\top V) \geq 0, & i \in \mathcal{V} \\ & \begin{bmatrix} \text{tr}(e_i e_i^* V) + \text{tr}(e_j e_j^* V) \\ 2\text{tr}(e_i e_j^\top V) \\ \text{tr}(e_i e_i^* V) - \text{tr}(e_j e_j^* V) \end{bmatrix} \in \mathbf{L}^3, & \forall\, 1 \leq i < j \leq n. \end{array} \tag{2.11}$$

By construction, it follows readily that the optimal value of the second-order cone relaxation (2.11) is a lower bound to the the optimal values of both the semidefinite relaxation (2.10) and the AC-OPF problem (2.6).

2.5 Exactness of Semidefinite Relaxations of AC-OPF

2.5.1 A Priori Sufficient Conditions for Exactness

The semidefinite relaxation of AC-OPF is said to be exact if and only if its optimal solution set contains a rank one matrix. Several papers [13, 49] have established sufficient conditions on the AC-OPF feasible region under which the semidefinite relaxation is exact for networks with acyclic topologies. Such topologies are typical in most electrical distribution systems. To state the results we first require some essential definitions.

Definition 2.5.1 (Linearly Separable)**.** *A set of complex numbers $x_1, \ldots, x_r \in \mathbf{C}$ is defined to be* linearly separable from the origin*, if there exists $0 \neq p \in \mathbf{C}$ such that $\mathrm{Re}(p^* x_k) \leq 0$ for all $k = 1, \ldots, r$.*

In other words, a set of complex numbers is linearly separable from the origin, if there exists a line through the origin of the complex plane such that all points in this set lie on one side of the line. The points may lie on the line, as the separation can be nonstrict. Figures 2.1(a)-(b) provide an example of a set of complex numbers which are linearly separable. The set of complex number in Figure 2.1(c) is not linearly separable. We now introduce *off-diagonal linear separability* of a semidefinite program, as defined in [13, 76].

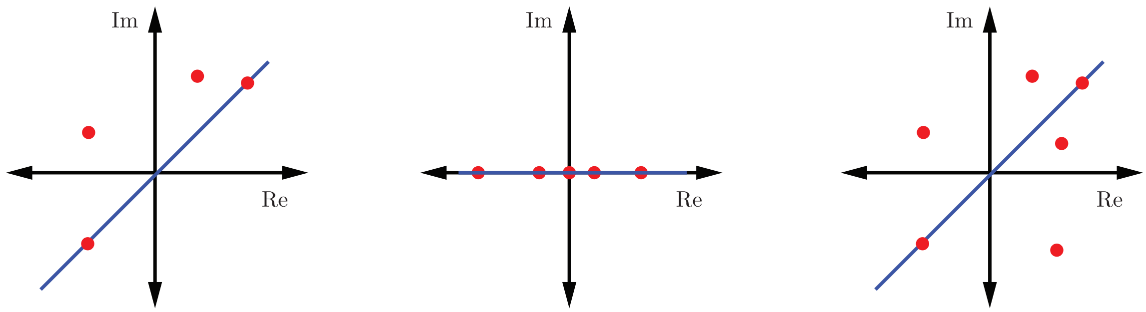


Figure 2.1: Figures (a) and (b) are examples of sets of complex numbers, which are linearly separable from the origin and (c) an example of a set that is not linearly separable from the origin.

Definition 2.5.2. *The semidefinite program* (2.10) *is defined to be off-diagonally linearly separable from the origin, if the set of complex numbers*

$$\{[A_0]_{ij}, [A_1]_{ij}, \dots, [A_m]_{ij}, \pm[A_{m+1}]_{ij}, \dots, \pm[A_{m+\ell}]_{ij}\} \tag{2.12}$$

is linearly separable from the origin for all $1 \leq i, j \leq n$ *and* $i \neq j$.

We are now in a position to state the following result from [13, 76]. It provides sufficient conditions, under which a rank one optimal solution to the semidefinite program (2.10) can be computed in polynomial time. The proof of Theorem 3.2.4 provides a simple procedure for constructing such a solution and we include it for completeness. Recall that $\mathcal{G}$ denotes the graph of the power network (cf. Section 2.2).

Theorem 2.5.3. *If* $\mathcal{G}$ *is acyclic and the semidefinite program* (2.10) *is off-diagonally linearly separable from the origin, then an optimal solution to* (2.10)*, which has rank one can be computed in polynomial time.*

Proof. The proof provides a procedure for constructing an optimal solution to the semidefinite relaxation that has rank one for an optimal solution of high rank. Without loss of generality, we assume throughout that the graph $\mathcal{G}$ is a connected acyclic graph, that is, it is a tree. Let

$$\mathcal{A} \coloneqq \{A_0, A_1, \ldots, A_m, \pm A_{m+1}, \ldots, \pm A_{m+\ell}\}$$

The proof proceeds in two steps. First, a matrix $W \in \mathbf{H}^n$ is constructed from the optimal solution $V^\star$ to the semidefinite relaxation (2.10), which satisfies $W_{ii}W_{jj} = |W_{jk}|^2$ for all $(i,j) \in \mathcal{E}$ and $\text{tr}(A_i W) \leq \text{tr}(A_i V^\star)$, for all $i = 0, \ldots, m+\ell$. Then, a vector $v \in \mathbf{C}^n$ is constructed from W, which satisfies

$$v^* A v = \text{tr}(AW) \leq \text{tr}(AV^\star), \qquad \text{for all } A \in \mathcal{A} \tag{2.13}$$

It is straightforward to see that above inequalities imply that v must be an optimal solution to the the nonconvex QCQP (2.9).

(i) *Constructing W from $V^\star$:* For each $i = 1, \ldots, n$, let $[W]_{ii} \coloneqq [V^\star]_{ii}$ and for each $(i,j) \in \mathcal{E}$, let

$$[W]_{ji} \coloneqq [V^\star]_{ji} + y_{ji} \exp^{\mathbf{i}(-\pi/2+\vartheta_{ji})},$$

where y_{ji} is a positive scalar to be specified below and ϑ_{ji} is an angle satisfying

$$\vartheta_{ji} \leq \measuredangle[A]_{ji} \leq \vartheta_{ji} + \pi, \quad \forall\, A \in \mathcal{A}. \tag{2.14}$$

Such an angle is guaranteed to exist since for each $(i,j) \in \mathcal{E}$, the set of complex numbers (2.12) is assumed to be linearly separable from the origin. Moreover, since W is a Hermitian matrix, it follows readily that $\theta_{ij} = (\pi - \theta_{ji}) \bmod 2\pi$. [3] All off-diagonal entries of W which are not in the set of edge $\mathcal{E}$ of $\mathcal{G}$ are left unspecified for the moment. Next, we argue that $\text{tr}(AW) \leq \text{tr}(AV^*)$, for all $A \in \mathcal{A}$, as long as

[3] $x = y \bmod z$ if and only if $(x-y) = kz$ for some integer k.

$y_{ij} \geq 0$. In particular, note that

$$\begin{aligned}
\text{tr}\,(A(W - X^\star)) &= \sum_{i=1}^{n} [A]_{ii}\big([W]_{ii} - [V^\star]_{ii}\big) + \sum_{(i,j)\in\mathcal{E}} [A]_{ij}\big([W]_{ji} - [V^\star]_{ji}\big) \\
&= \sum_{\substack{(i,j)\in\mathcal{E} \\ i<j}} 2\,\mathrm{Re}\Big\{[A]_{ij}\big([W]_{ji} - [V^\star]_{ji}\big)\Big\} \\
&= \sum_{\substack{(i,j)\in\mathcal{E} \\ i<j}} 2\big|[A]_{ij}\big| y_{ji} \cos(\measuredangle[A]_{ij} - \pi/2 + \vartheta_{ji}),
\end{aligned}$$

where all equalities follow from the definition of W. In addition, it follows readily by equation (2.14) that the argument of the cosine in the above expression lies in the interval $[\pi/2, 3\pi/2]$ and therefore, $\cos(\measuredangle[A]_{ij} - \pi/2 + \vartheta_{ji}) \leq 0$. Therefore, if $y_{ji} \geq 0$, we must have

$$\text{tr}\,(A(W - V^\star)) \leq 0. \tag{2.15}$$

We choose y_{ji} so that we have $W_{ii}W_{jj} = |W_{ji}|^2$, or equivalently

$$[V^\star]_{ii}[V^\star]_{jj} \;=\; \left|[V^\star]_{ji} + y_{ji}e^{\mathbf{i}(-\pi/2+\vartheta_{ji})}\right|.$$

This is a quadratic relation in y_{ji}, which admits a closed form solution given by $y_{ji} = \sqrt{b^2 + c} - b$, where

$$b := \mathrm{Re}\left\{[V^\star]_{ji}e^{\mathbf{i}(\pi/2-\vartheta_{ji})}\right\}, \qquad c := [V^\star]_{ii}[V^\star]_{jj} - \left|[V^\star]_{ji}\right|^2.$$

Since $V^\star$ is positive semidefinite, the 2×2 principal minor corresponding to its i^{th} and j^{th} columns is also positive semidefinite. Therefore, $c \geq 0$, which in turn implies that $y_{ji} \geq 0$.

(i) *Constructing a rank-one matrix from W:* We construct a vector $v \in \mathbf{C}^n$ satisfying (2.13). The rank one matrix is obtained by taking the outer product of v with itself.

For each $i = 1, \ldots, n$, let $v_i := \sqrt{W_{ii}}$ and set $\measuredangle v_1 := 0$. And for each node $j \in \{2, \ldots, n\}$, let $(\ell_0, \ell_1), (\ell_1, \ell_2), \ldots, (\ell_k, \ell_j)$ be the unique path from node 1 to node j in the tree, where $\ell_0 = 1$ and $\ell_j = j$. We define

$$\measuredangle v_j := -\sum_{k=0} \measuredangle [W]_{\ell_k \ell_{k+1}}.$$

Note that for $(i,j) \in \mathcal{E}$, we have $W_{ji} = \measuredangle v_j - \measuredangle v_i$. Equality (2.13) is satisfied since

$$\begin{aligned} v^* A v - \operatorname{tr}(AW) &= \sum_{i=1}^{n} [A]_{ii} (|v_i|^2 - W_{ii}) + \sum_{\substack{i \neq j \\ 1 \leq i,j \leq n}} [A]_{ij} (v_i^* v_j - W_{ji}) \\ &= \sum_{(i,j) \in \mathcal{E}} 2\mathrm{Re} \left\{ [A]_{ij} (v_i^* v_j - W_{ji}) \right\} \\ &= \sum_{(i,j) \in \mathcal{E}} 2\mathrm{Re} \left\{ [A]_{ij} \left(|v_i||v_j| e^{\mathbf{i}(\measuredangle v_j - \measuredangle v_i)} - W_{ji} \right) \right\} \\ &= \sum_{(i,j) \in \mathcal{E}} 2\mathrm{Re} \left\{ [A]_{ij} \left(\sqrt{W_{ii} W_{jj}} e^{\mathbf{i} \measuredangle W_{ji}} - W_{ji} \right) \right\} \\ &= 0. \end{aligned} \tag{2.16}$$

It follows readily from expressions (2.15) and (2.16) that the rank-one matrix vv^* is an optimal solution for the semidefinite relaxation (2.10).

■

We now discuss the limitations of Theorem 3.2.4 to the semidefinite relaxation of AC-OPF. To facilitate our discussion, we restate Theorem 3.2.4 using the definition of the matrices $\Phi_i, \Psi_i,\ i \in \mathcal{V}$ in (2.8) and $P_{ij},\ (i,j) \in \mathcal{E}$ in (2.3). We have the following Corollary.

Corollary 2.5.4. *If the power network is characterized by an acyclic graph and the matrices*

$$\pm\Phi_i, \pm\Psi_i, P_{ij}, \quad i \in \mathcal{V}, \ \ (i,j) \in \mathcal{E}, \qquad \Phi_0 := \sum_{i \in \mathcal{V}_g} \alpha_i \Phi_i$$

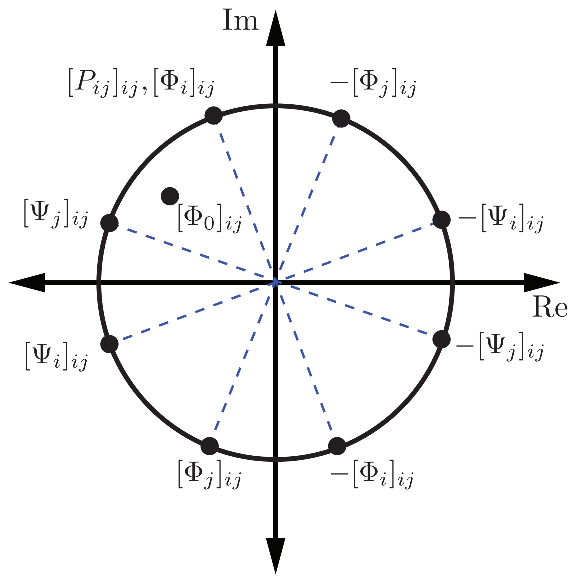


Figure 2.2: This figure from [49] depicts the linear separability condition on a line $(i, j) \in \mathcal{E}$. The quantities $([\Phi_i]_{ij}, [\Phi_j]_{ij}, [\Psi_i]_{ij}, [\Psi_j]_{ij})$ correspond to upper bounds on the active and reactive power injections. Their negative quantities correspond to lower bounds. The linear separability assumption is satisfied if there is a line through the origin such that the quantities corresponding to finite upper and lower bounds on the active and reactive power injections lie on one side of the line. The figure illustrates that the linear separability condition cannot be satisfied if there is a line where both real and reactive power injections at both ends of the line are both upper and lower bounded by finite numbers. Power networks with at least one load bus, i.e., a bus $i \in \mathcal{V} \setminus \mathcal{V}_g$ naturally violate the linear separability condition.

defined in (2.3) *and* (2.8) *are off-diagonally linearly separable from the origin, then an optimal solution to* (2.10)*, which has rank equal to one can be computed in polynomial time.*

We refer the reader to Figure 2.2, which is taken from [49] for a graphical illustration of the limitations of the Theorem. The linear separability condition cannot be satisfied if there is a line where both real and reactive power injections at both ends of the line are both upper and lower bounded (by finite numbers). Therefore, Theorem 3.2.4 cannot be applied to power networks with at least one load bus, i.e., a bus $i \in \mathcal{V} \setminus \mathcal{V}_g$. Consider now the case where a transmission line $(i, j) \in \mathcal{E}$ connects two generator buses, i.e., $i, j \in \mathcal{V}_g$.

The real and reactive power injections at buses i, j are required to satisfy

$$v^*\Phi_k v \le \mathrm{Re}\{\overline{g}_k - d_k\}, \qquad v^*\Psi_k v \le \mathrm{Im}\{\overline{g}_k - d_k\}, \tag{2.17}$$

$$v^*(-\Phi_k)v \le -\mathrm{Re}\{\underline{g}_k - d_k\}, \qquad v^*(-\Psi_k)v \le -\mathrm{Im}\{\underline{g}_k - d_k\}, \tag{2.18}$$

where $k = i, j$ and the matrices Φ_k, Ψ_k are defined in (2.8). One can verify that $[\Phi_i]_{ij} = ([\Phi_j]_{ij})^*$ and $[\Psi_i]_{ij} = ([\Psi_j]_{ij})^*$ (cf. Figure 2.2). Therefore, the linear separability assumption holds if and and only if either the right hand side of constraints (2.17) is unbounded i.e., $\overline{g}_k = +\infty + \mathbf{i}\infty$, for $k = i, j$ or if the right hand side of constraints (2.18) is unbounded, i.e., $\underline{g}_k = -\infty - \mathbf{i}\infty$ for $k = i, j$.

Remark 2 (Load Oversatisfaction). *A number of sufficient conditions in the literature guaranteeing exactness of solutions to semidefinite relaxations of OPF over radial networks rely on the so-called assumption of load-oversatisfaction. In part, the assumption of load-oversatisfaction amounts to relaxing the power balance equations (at load buses) to inequalities, where the complex power delivered to each node is allowed to exceed the power demanded. The load oversatisfaction assumption is a special case of the linear separability condition, where the line is always taken to be equal to the vertical line passing through the origin (i.e., the imaginary axis).*

2.5.2 A Posteriori Sufficient Conditions for Inexactness

While the importance of exactness results over networks with acyclic topologies is self-evident, there remains an incomplete understanding of the behavior of semidefinite relaxations for more general network structures and AC-OPF feasible regions. In practice it has been observed that many instances of AC-OPF yield semidefinite relaxations with optimal solutions of high rank – even though rank one optimal solutions may exist. In this section, we provide an a posteriori method for verifying that the semidefinite relaxation is inexact. In particular, we exploit dual nondegeneracy of semidefinite programs – as defined by Alizadeh et al. in [1] – to establish uniqueness of primal optimal solutions of the semidefinite relaxation. If the optimal solution of the semidefinite relaxation has rank greater than one and it is the unique solution, then the semidefinite relaxation must be inexact. We refer the reader to Appendix B.4 for a brief introduction on constraint nondegeneracy for semidefinite programs.

To state our results, let us consider the dual problem of the semidefinite relaxation of AC-OPF (2.10), which is given by

$$\begin{aligned} \underset{y\in\mathbf{R}^{m+\ell},\ Z\in\mathbf{H}^n}{\text{maximize}} \quad & -b^\top y \\ \text{subject to} \quad & Z = A_0 + \sum_{i=1}^{m+\ell} y_i A_i, \\ & y_i \geq 0, \quad i = m+1, \dots, m+\ell \\ & Z \succeq 0 \end{aligned} \tag{2.19}$$

where $b := [b_1, \dots, b_{m+\ell}] \in \mathbf{R}^{m+\ell}$. For any feasible matrix V of the primal problem (2.10), let

$$\mathcal{I}(V) := \{i \mid \mathrm{tr}(A_i V) = b_i,\ i = 1, \dots, m\}$$

be the set of inequality constraints that are active at V.

We make the following assumptions about the primal-dual pair (2.10) – (2.19), which apply throughout this section.

Assumption 2.5.5. There exists a matrix $V \in \mathbf{H}^n$, which is feasible for (2.10) and satisfies $V \succ 0$ and $\mathrm{tr}(A_i V) < b_i$ for all $i = 1, \dots, m$. Moreover, there exists a vector $y \in \mathbf{R}^{m+\ell}$, which is feasible for (2.19) such that $Z \succ 0$ and $y_i > 0$ for all $i = 1, \dots, m$.

Assumption 3.2.1 is a Slater Condition and it guarantees that strong duality holds. This, in turn, implies a complementarity condition [1]. Namely, for any pair of primal-dual optimal solutions $(V^\star, y^\star, Z^\star)$, it holds that $\operatorname{rank}(V^\star)+\operatorname{rank}(Z^\star) \leq n$. A primal-dual optimal solution pair $(V^\star, y^\star, Z^\star)$ is said to satisfy *strict complementarity* if $\operatorname{rank}(V^\star) + \operatorname{rank}(Z^\star) = n$.

Assumption 2.5.6. Strict complementarity holds between any pair of primal-dual optimal solutions.

Assumption 2.5.7. The matrices in the set $\{A_i \mid i = m+1, \ldots, m+\ell\}$ are linearly independent.

The linear independence assumption is without loss of generality. If the matrices A_i, $i = m+1, \ldots, m+\ell$ are linearly dependent, one can choose a basis of, say $p < \ell$, matrices from $\{A_i \mid i = m+1, \ldots, m+\ell\}$ and remove the other $\ell - p$ equality constraints to establish an equivalent problem in which the assumption holds.

Theorem 2.5.8 (Inexactness of Semidefinite Relaxation). *Let $(V^\star, y^\star, Z^\star)$ be a set of primal-dual optimal solutions to (2.6)-(2.19) and suppose that $\operatorname{rank}(Z^\star) = r$. Let $Z^\star = Q\Sigma Q^\top$ be an eigenvalue decomposition of of $Z^\star$, where $\Sigma = \operatorname{diag}(0, \ldots, \sigma_{n-r+1}, \sigma_n) \in \mathbf{R}^{n\times n}$ and $Q \in \mathbf{C}^{n\times n}$. Partition Q as $Q = [Q_1, Q_2]$, where $Q_1 \in \mathbf{C}^{n-r}$ and $Q_2 \in \mathbf{C}^{n\times r}$, and define the matrices*

$$B_k := Q_1^* A_k Q_1, \quad \textit{for all } i \in \mathcal{I}(V^\star) \cup \{m+1, \ldots, m+\ell\}.$$

If the matrices $B_i, i \in \mathcal{I}(V^\star) \cup \{m+1, \ldots, m+\ell\}$ span $\mathbf{H}^{n-r}$, then $V^\star$ is the unique solution to the semidefinite relaxation (2.10). In addition, if $\operatorname{rank}(V^\star) > 1$, the semidefinite relaxation of AC-OPF is inexact.

We summarize, in Table 2.1, several AC-OPF test cases from the literature [43, 96, 86] whose corresponding semidefinite relaxations are guaranteed to be inexact according to

Power System (n)	Reference	$\mathrm{rank}(V^\star)$	ℓ	$\lvert\mathcal{I}(V^\star)\rvert$
3	[43]	2	1	4
5	[86]	2	6	3
39[4]	[96]	2	58	14
118	[96]	2	128	73

Table 2.1: Power system examples with semidefinite relaxations having unique high rank solutions.

Theorem 2.5.8. Namely, their semidefinite relaxations admit *unique high rank optimal solutions*. Other examples in the literature that yield high rank optimal solutions include the 9 and 30 bus networks in [96]. For these AC-OPF problems, however, the solution to the semidefinite relaxation is dual degenerate. By adding a small resistance (e.g. 10^{-5} ohms) to a subset of the lines with zero resistance, we obtain a dual nondegenerate optimal solution and a (unique) rank one primal optimal solution whose cost is within 0.002% of the optimal value in the degenerate case.

2.6 Nonconvex Optimization Methods

Since its original formulation by Carpentier [17] in 1962, a variety of techniques from mathematical programming have been proposed to solve the AC-OPF problem (2.6). These include – but are not limited to – sequential quadratic programming methods [80, 16], augmented Lagrangian methods [70], primal-dual interior point methods [84, 57, 58], predictor-corrector methods [89, 90, 91, 19], and trust-region methods [77].

In section 2.6.1 we focus on a primal-dual interior point algorithm based on a barrier function method, which is incorporated in the software package MATPOWER [96]. In addition,

[4]A small resistance of 10^{-5} was added to all lines having zero resistance.

we discuss a consensus alternating direction method of multipliers that was recently proposed by Huang and Sidiropoulos [32] for general nonconvex quadratically constrained quadratic programs.

2.6.1 Primal-Dual Interior Point Algorithms

The treatment of nonconvexities in AC-OPF has traditionally relied on the use of local methods for constrained optimization. In this section, we review the primal-dual interior point algorithm MIPS, which is incorporated in the software package MATPOWER [96]. MIPS is an interior point algorithm based on a barrier function method. To apply the algorithm, a set of slack variables is first introduced to transform the inequality constraints in AC-OPF into a set of equality constraints. A log barrier penalty term is then added to the objective function to enforce nonnegativity of the slack variables. This gives rise to an optimization problem, which only involves equality constraints. A locally optimal solution is obtained from Newton's method, which is used to solve the first-order necessary (KKT) conditions for local optimality.

Let $z_1, \ldots, z_m \in \mathbf{R}$ be slack variables for the inequality constraints of the QCQP formulation of the AC-OPF problem (2.9) and consider the following optimization problem,

$$\begin{array}{lll} \underset{v \in \mathbf{C}^n}{\text{minimize}} & v^* A_0 v - \gamma \sum_{i=1}^{m} \log(z_i) & \\ \text{subject to} & v^* A_i v + z_i = b_i, & i = 1, \ldots, m \\ & v^* A_i v = b_i, & i = m+1, \ldots, m+\ell, \end{array} \tag{2.20}$$

where $\gamma > 0$ is a parameter that sets the accuracy of the approximation. In particular, as γ decreases, the approximation becomes more accurate. For a given value of γ, the

Algorithm 2.2: Primal-Dual Interior Point Algorithm MIPS

Given an initial condition $(\Delta x^0, \Delta z^0, \Delta\lambda^0)$, a stopping tolerance $\varepsilon > 0$, a scalar $\zeta \in (0,1)$, and maximum number of iterations $\overline{k}$

Initialize $k = 0$

Repeat

1. *Compute.* $v^{k+1} = v^k + \alpha_p^k \Delta v^k$
2. *Compute.* $z^{k+1} = z^k + \alpha_p \Delta z^k$
3. *Compute.* $\lambda^{k+1} = \alpha_d^k \Delta\lambda^k$
4. *Update.* $\alpha_p^{k+1} = \min\left\{\zeta \min_{i:\Delta z_i^{k+1}<0}\left\{\dfrac{-z_i^{k+1}}{\Delta z_i^{k+1}}\right\}, 1\right\}$,
5. *Update.* $\alpha_d^{k+1} = \min\left\{\zeta \min_{1\le i\le m:\Delta\lambda_i^{k+1}<0}\left\{\dfrac{-\lambda_i^{k+1}}{\Delta\lambda_i^{k+1}}\right\}, 1\right\}$,
6. *Update.* $\gamma^{k+1} = (0.1/m)\displaystyle\sum_{i=1}^{m} z_i^{k+1}\lambda_i^{k+1}$
7. *Update.* $k = k+1$

Until. $|v^k - v^{k-1}| < \varepsilon$

Output. v^k

Table 2.2: Primal-dual Interior point algorithm MIPS incorporated in the software package MATPOWER for the AC-OPF problem

Lagrangian function $L_\gamma : \mathbf{C}^n \times \mathbf{R}^m \times \mathbf{R}^{m+\ell} \to \mathbf{R}$ for problem (2.20) is given by

$$L_\gamma(v, z, \lambda) = v^* \left(A_0 + \sum_{i=1}^{m+\ell} \lambda_i A_i\right) v - \gamma \sum_{i=1}^{m} \log(z_i) + \sum_{i=1}^{m} \lambda_i z_i - b^\top \lambda,$$

where $b = [b_1, \ldots, b_{m+\ell}]^\top$. The first-order necessary conditions for local optimality require that the gradient of the Lagrangian function L_γ with respect to the variables (v, z, λ) be zero, i.e.,

$$
\begin{aligned}
\frac{\partial L_\gamma(v,z,\lambda)}{\partial v} &= \left(A_0 + \sum_{i=1}^{m+\ell} \lambda_i A_i\right) v = 0, \\
\frac{\partial L_\gamma(v,z,\lambda)}{\partial z_i} &= -\frac{\gamma}{z_i} + \lambda_i = 0, \qquad\qquad i = 1,\ldots,m \\
\frac{\partial L_\gamma(v,z,\lambda)}{\partial \lambda} &= h(v) + z - b = 0,
\end{aligned}
$$

where $h : \mathbf{C}^n \to \mathbf{R}^m$ is a vector-valued function given by $h(v) := [v^* A_1 v \ \cdots \ v^* A_{m+\ell} v]$. The above system of equations is solved by Netwon's Method. The Newton step $(\Delta v, \Delta z, \Delta\lambda)$ is a solution to the linear system

$$
N(v,z,\lambda) \begin{bmatrix} \Delta v \\ \Delta z \\ \Delta \lambda \end{bmatrix} = -F_\gamma(v,z,\lambda)
$$

where,

$$
N(v,z,\lambda) := \begin{bmatrix} A_0 + \sum_{i=1}^m \lambda_i A_i & 0 & H(v) \\ 0 & \mathrm{diag}(\lambda_1,\ldots,\lambda_m) & \begin{bmatrix} \mathrm{diag}(z) & 0_\ell \end{bmatrix} \\ H(v)^\top & \begin{bmatrix} I_m & 0_{m\times\ell} \end{bmatrix}^\top & 0 \end{bmatrix}, \tag{2.21}
$$

and

$$
H(v) := \begin{bmatrix} A_1 v, & \ldots & A_{m+\ell} v \end{bmatrix} \quad \text{and} \quad F_\gamma(v,z,\lambda) := \begin{bmatrix} \left(A_0 + \sum_{i=1}^{m+\ell} \lambda_i A_i\right) v \\ \mathrm{diag}(\lambda_1,\ldots,\lambda_m) z - \gamma \mathbf{1}_m \\ h(v) + z - b \end{bmatrix}.
$$

Here, $\mathbf{1}_m$ denotes the m-dimensional vector of all ones. The primal-dual interior point algorithm MIPS is given in Table 2.2. The superscript k indicates the value of the corresponding variable at the k^{th} iteration of the algorithm. The scalars α_p and α_d are step sizes

Algorithm: Consensus ADMM

Given an initial condition $(v^0, z_1^0, \ldots, z_{m+\ell}^0, u^0)$, a stopping tolerance $\varepsilon > 0$, and maximum number of iterations $\overline{k}$

Initialize $k = 0$

Repeat

1. *Compute.* $v^{k+1} = \rho(A_0 + (m+\ell)\rho I)^{-1} \sum_{i=1}^{m} (z_i^k + u_i^k)$
2. For each $i = 1, \ldots, m$:
   (i) *Compute.* $z_i^{k+1} = \operatorname*{argmin}_{z_i \in \mathbf{C}^n} \quad \|z_i - v^{k+1} + u_i^k\|_2^2 \qquad \text{subject to} \qquad z_i^* A_i z_i \leq b_i,$
   (ii) *Compute.* $u_i^{k+1} = u_i^k + z_i^{k+1} - v^{k+1}$
3. For each $i = m+1, \ldots, m+\ell$:
   (i) *Compute.* $z_i^{k+1} = \operatorname*{argmin}_{z_i \in \mathbf{C}^n} \quad \|z_i - v^{k+1} + u_i^k\|_2^2 \qquad \text{subject to} \qquad z_i^* A_i z_i = b_i,$
   (ii) *Compute.* $u_i^{k+1} = u_i^k + z_i^{k+1} - v^{k+1}$
4. *Update.* $k = k + 1$

Until $|v^{k+1} - v^k| < \varepsilon,$ and $|z^i - v_k| < \varepsilon,$ for all $i = 1, \ldots, m+\ell$

Output $(v^k, z_1^k, \ldots, z_{m+\ell}^k, u^k)$

Table 2.3: Consensus ADMM Algorithm for the QCQP Formulation of the AC-OPF problem.

which are chosen to maintain strict feasibility of the solution at each step. The parameter ζ is a positive constant whose value is value slightly less than one. In MIPS it is chosen to be equal to 0.99995. MIPS uses the rule at step 6 of the algorithm to update γ at each iteration.

2.6.2 Consensus Alternating Direction Method of Multipliers

In this section, we introduce a consensus alternating direction method of multipliers (ADMM) approach for the AC-OPF problem (2.9). This approach was first proposed by Huang and Sidiropoulos [32] for general nonconvex quadratically constrained quadratic programs. We refer the reader to Appendix B.2 for a brief introduction on the consensus form of ADMM. The predominant advantage that this approach offers is that it decomposes the original AC-OPF problem (2.9) into $m + \ell$ subproblems, each of which can be solved optimally, albeit nonconvex. However, due to the nonconvexity of AC-OPF, this algorithm is not guaranteed to converge to a KKT point (cf. Appendix B.1) of (2.9).

In order to transform AC-OPF into a form suitable for the consensus ADMM algorithm, we introduce one decision variable for each constraint of (2.9). More precisely, for each $i = 1, \ldots, m + \ell$, let $z_i \in \mathbf{C}^n$ and consider the following optimization problem in $m + \ell + 1$ variables

$$
\begin{array}{lll}
\underset{v, z_1, \ldots, z_{m+\ell}}{\text{minimize}} & v^* A_0 v & \\
\text{subject to} & z_i^* A_i z_i \leq b_i, & i = 1, \ldots, m \\
& z_i^* A_i z_i = b_i, & i = m+1, \ldots, m+\ell, \\
& z_i = v & i = 1, \ldots, m+\ell.
\end{array}
\tag{2.22}
$$

Clearly, the projection of an optimal solution of (2.22) onto the set of v variables is an optimal solution to the AC-OPF problem. The corresponding consensus ADMM iterates, are described in Table 2.3. The superscript k indicates the value of the corresponding variable at the k^{th} iteration of the algorithm. We refer the reader to Appendix B.2 for a brief discussion on the consensus ADMM algorithm.

Some remarks regarding the ADMM algorithm are in order. First, each optimization problem in steps 2(i) and 3(i) of the ADMM algorithm in Table 2.3 is a nonconvex quadratically constrained quadratic programs with one constraint. This optimization problem can be solved in polynomial time through its semidefinite programming relaxation. We refer the reader to Corollary 4.2.3 in Appendix B.7 for the corresponding theoretical result. In addition, at each iteration step k, the variables $z_i^k,\ i = 1, \ldots, m + \ell$ can be updated in parallel. The following Theorem from [32] states that if the consensus ADMM algorithm (2.3) converges, then it converges to a KKT point of (2.9).

Theorem 2.6.1. *If*

$$\lim_{k\to\infty} (z_i^k - x^k) = 0, \qquad \textit{for all } i = 1, \ldots, m + \ell$$

and

$$\lim_{k\to\infty} (x^{k+1} - x^k) = 0,$$

then any limit point of the sequence $\{x^k\}$ *is a KKT point of* (2.9).

2.7 Conclusions

This chapter considered the AC optimal power flow (AC-OPF) problem, which can be formulated as a nonconvex quadratically constrained quadratic problem. This formulation gives rise to a second-order cone and a semidefinite relaxation for the AC-OPF problem. We presented an a priori sufficient condition which was developed in [13] stating that power systems with acyclic topologies satisfying a certain condition (linear separability) admit exact semidefinite relaxations. In addition, we developed an a posteriori sufficient condition to verify that the optimal solution face of the semidefinite relaxation does not contain a rank one matrix (i.e., the semidefinite relaxation is inexact). Namely, our sufficient condition can be used to verify whether an optimal solution to the semidefinite relaxation is the unique optimal solution. Using this sufficient condition, we identify a number of power system examples from the literature which yield inexact semidefinite relaxations. Finally, we introduce two nonconvex optimization approaches for obtaining locally optimal solutions to the AC-OPF problem. The first is based on a primal-dual interior point method and the second on the alternating direction method of multipliers.

CHAPTER 3

PERTURBATION METHODS FOR QUADRATICALLY CONSTRAINED QUADRATIC PROGRAMS

3.1 Introduction

A quadratically constrained quadratic program (QCQP) is a nonlinear optimization problem that can be expressed in the form

$$\begin{array}{ll} \underset{x \in \mathbf{C}^n}{\text{minimize}} & x^* A_0 x \\ \text{subject to} & x^* A_k x \le b_k, \quad \text{for all } k = 1, \ldots, m, \end{array} \tag{3.1}$$

where the scalars $b_1, \ldots, b_m$ and the Hermitian matrices $A_0, A_1, \ldots, A_m$ are the given problem data. The variable $x \in \mathbf{C}^n$ is a complex vector and x^* is its complex conjugate transpose. As such, QCQPs are nonlinear optimization problems in which both the objective and the constraints are described by quadratic functions.

Quadratically constrained quadratic programs are ubiquitous in many branches of applied mathematics and engineering. Many engineering problems such as sensor network localization [75], MIMO detection [74, 39], multicast downlink transmit beamforming [24], and optimal power flow [13, 47, 76], can be formulated as QCQPs. Moreover, QCQPs find a wide applicability in the domains of combinatorial optimization and graph theory, as 0-1 integer programming problems can be equivalently reformulated as QCQPs. These include the max-cut problem, the maximum 2-satisfiability problem (MAX 2-SAT) [26], and the quadratic knapsack problem [29], to name a few.

In their most general form, QCQPs are nonconvex optimization problems that are NP-hard.

This has motivated the design and study of approximation algorithms capable of efficiently computing approximate solutions to QCQPs with theoretical guarantees on performance. One such technique is the *semidefinite relaxation*, which entails the relaxation of the QCQP to a linear program over the cone of positive semidefinite matrices [51].

The semidefinite relaxation technique has sparked two lines of research: one direction focusing on the design of approximation algorithms for QCQPs based on semidefinite programming [26], [8, 60, 61, 92, 93, 28] and another exploring the characterization of classes of QCQPs for which an optimal solution can be computed in polynomial time from an optimal solution to its semidefinite relaxation [13, 37, 95].

In the former direction, the majority of the work has considered special classes of QCQPs. For example, Nemirovski et al. [60] provide an $O(\log m)$-approximation algorithm for QCQPs defined over nonempty convex compact feasible sets. In particular, the data $A_1, \ldots, A_m$ are arbitrary positive semidefinite matrices whose sum is positive definite. Building on this work, He et al. [28] show that the same approximation guarantee holds when exactly one of the matrices $A_1, \ldots, A_m$ is indefinite. The authors in [28] also provide a data dependent approximation ratio for the class of QCQPs with an arbitrary number of indefinite quadratic constraints, assuming that $b_k > 0$ for all $k = 1, \ldots, m$. The authors do not, however, offer a systematic procedure guaranteed to recover a feasible solution to the QCQP with indefinite constraints.

In the latter direction, Kim et al. [37] prove that real QCQPs having problem data $A_0, A_1, \ldots, A_m$ that respect certain conditions on sign definiteness admit semidefinite relaxations whose optimal solutions can be efficiently mapped to a globally optimal solution of the corresponding QCQP. More recently, several authors [13, 76] have generalized this result to the case of complex QCQPs with acyclic graph structure.

Contribution: In this chapter, we consider complex QCQPs having arbitrary graph structure and problem data that allow for indefinite matrices $A_0, A_1, \ldots, A_m$ and negative scalars $b_1, \ldots, b_m$. The crux of our approach centers on the application of a *structured perturbation* to the problem data to yield a perturbed QCQP whose feasible set is a *nonempty* subset of the original feasible set, whose constraints satisfy the technical conditions in [13, 76], and whose collective sparsity pattern defines an *acyclic graph*. By relaxing the perturbed QCQP to a semidefinite program, one can compute in polynomial time a feasible solution to the perturbed QCQP and, hence, the original QCQP. We refer to this procedure as the *acyclic semidefinite approximation* of the QCQP. The challenge lies in designing the perturbation to ensure both nonemptiness of the perturbed feasible set and acceptable bounds on the performance of the suboptimal solution obtained. Leveraging on the notion of the *distance to infeasibility* of a conic linear system – as defined by Renegar in [67] – we give a sufficient condition under which a perturbation is guaranteed to yield a nonempty feasible set and provide a bound on the performance of optimal solutions to the perturbed problem. The performance guarantee depends on both the original problem data and the size of the perturbation.

Organization: The remainder of the paper is organized as follows. In Section 3.2, we formulate the semidefinite relaxation, present assumptions, and provide a definition of acyclic semidefinite approximations. Section 3.3 contains our main results, which characterize the performance of acyclic semidefinite approximations. Conclusions and directions for future research are given in Section 3.4.

Notation: Let $\mathbf{R}$ be the field of real numbers and $\mathbf{C}$ the field of complex numbers. For

$z \in \mathbf{C}$, let $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$ be the real and imaginary parts of z, respectively. Let $\mathbf{R}^n$ be the n-dimensional real vector space and $\mathbf{H}^n$ be the space of $n \times n$ Hermitian matrices. Given a matrix A, let $[A]_{ij}$ be its (i, j) entry. And denote by $A^\top$ and A^* the transpose and complex conjugate transpose of A, respectively. For a matrix $A \in \mathbf{H}^n$, the notation $A \succeq 0$ ($A \succ 0$) means that A is positive semidefinite (positive definite). Endow $\mathbf{R}^n$ with the inner product $x^\top y$ for all $x, y \in \mathbf{R}^n$. The induced norm is denoted by $\|x\|_2 := \sqrt{x^\top x}$. Also, endow $\mathbf{H}^n$ with the trace inner product $\mathrm{tr}(X^* Y)$ for all $X, Y \in \mathbf{H}^n$. The induced norm is denoted by $\|X\|_F := \sqrt{\mathrm{tr}(X^* X)}$. For a linear map $\mathcal{A} : \mathbf{H}^n \to \mathbf{R}^m$, let $\|\mathcal{A}\| := \max\{\|\mathcal{A}(X)\|_2 \mid X \in \mathbf{H}^n \text{ with } \|X\|_F \leq 1\}$ denote its operator norm. The adjoint of $\mathcal{A}$ is denoted by $\mathcal{A}^*$.

3.2 Preliminaries

3.2.1 Semidefinite Relaxation

Central to our analysis is the *semidefinite relaxation* of the QCQP in (3.1). Its derivation entails the exact reformulation of the QCQP as a semidefinite program, whose feasible set is restricted to the space of rank one matrices. The semidefinite relaxation is obtained by removing the rank constraint. It is defined as:

$$\underset{X \in \mathbf{H}^n}{\text{minimize}} \quad \mathrm{tr}(A_0 X) \quad \text{s.t.} \quad \mathcal{A}(X) \leq b, \quad X \succeq 0, \tag{3.2}$$

where $\mathcal{A} : \mathbf{H}^n \to \mathbf{R}^m$ is the linear map defined by

$$\mathcal{A}(X) := [\mathrm{tr}(A_1 X), \ldots, \mathrm{tr}(A_m X)]^\top.$$

The dual of problem (3.2) is

$$\underset{y \in \mathbf{R}^m}{\text{maximize}} \quad -b^\top y \quad \text{s.t.} \quad A_0 + \mathcal{A}^*(y) \succeq 0, \;\; y \geq 0, \tag{3.3}$$

where $\mathcal{A}^* : \mathbb{R}^m \to \mathbf{H}^n$ denotes the adjoint of the linear map $\mathcal{A}$. It is given by $\mathcal{A}^*(y) = \sum_{k=1}^m y_k A_k$. The primal-dual pair of semidefinite programs (3.2)-(3.3) is fully specified by its *data*, which we denote by $d := (\mathcal{A}, b, A_0)$. Henceforth, we will succinctly refer to the primal-dual pair of programs (3.2)-(3.3) as the *semidefinite program* d.

The set of primal feasible solutions is defined by $\mathcal{P}(d) := \{X \succeq 0 \mid \mathcal{A}(X) \leq b\}$ and the set of dual feasible solutions by $\mathcal{D}(d) := \{y \geq 0 \mid A_0 + \mathcal{A}^*(y) \succeq 0\}$. We make the following assumption, which applies throughout the paper.

Assumption 3.2.1. The primal feasible set $\mathcal{P}(d)$ is nonempty and there exists a dual feasible solution $y \in \mathcal{D}(d)$ such that $y > 0$ and $A_0 + \mathcal{A}^*(y) \succ 0$. ■

Assumption 3.2.1 (a Slater condition) guarantees strong duality to hold. The importance of this assumption is made apparent in the proof of Theorem 3.3.3. Operating under only Assumption 3.2.1, the semidefinite program d may not admit rank one primal optimal solutions. The nonexistence of rank one primal optimal solutions does not, however, preclude the existence of rank one primal feasible solutions that are *near optimal*. In the following section, we explore the extent to which a *structured perturbation* might be applied to the problem data in order to force a primal optimal solution to the perturbed semidefinite program that is rank one and both feasible and nearly optimal for the original semidefinite program.

3.2.2 Acyclic Semidefinite Approximations

Our approach is motivated by a result from [13, 76], which characterizes a family of QCQPs that admit semidefinite relaxations having rank one optimal solutions that can be computed in polynomial time. We first require two essential definitions.

Definition 3.2.2. *Define the* graph of a semidefinite program *specified by data* d *as* $\mathcal{G}(d)$*, where* $\mathcal{G}(d) = (\mathcal{V}, \mathcal{E})$ *is a* simple graph *having a vertex set* $\mathcal{V} = \{1, \ldots, n\}$ *and edge set*

$$\mathcal{E} = \{(i,j) \mid i \neq j,\ \exists\, k = 0, 1, \ldots, m,\ [A_k]_{ij} \neq 0\}.$$

Remark 3. *Essentially, the graph* $\mathcal{G}(d)$ *of a semidefinite program* d *is an undirected graph, whose edges reflect the collective sparsity pattern of the matrices* $A_0, A_1, \ldots, A_m$*.*

We now introduce *off-diagonal linear separability* of a semidefinite program, as defined in [13, 76]. First, a set of complex numbers $x_0, \ldots, x_m \in \mathbf{C}$ is defined to be *linearly separable from the origin*, if there exists $0 \neq p \in \mathbf{C}$ such that $\mathrm{Re}(p^* x_k) \leq 0$ for all $k = 0, \ldots, m$. In other words, a set of complex numbers is linearly separable from the origin, if there exists a line through the origin of the complex plane such that all points in this set lie on one side of the line. The points may lie on the line, as the separation can be nonstrict.

Definition 3.2.3. *A semidefinite program* d *is defined to be off-diagonally linearly separable from the origin, if the set of complex numbers* $\{[A_0]_{ij}, [A_1]_{ij}, \ldots, [A_m]_{ij}\}$ *is linearly separable from the origin for all* $1 \leq i, j \leq n$ *and* $i \neq j$*.* ■

We are now in a position to state the following result from [13, 76]. It provides sufficient conditions on the data of semidefinite program d, under which a rank one optimal solution can be computed in polynomial time. Suppose that Assumption 3.2.1 holds.

Theorem 3.2.4. *Consider a semidefinite program d. If $G(d)$* is acyclic *and d* is off-diagonally linearly separable from the origin, *then a rank one primal optimal solution can be computed in polynomial time.* ■

The basic implication of Theorem 3.2.4 is that complex QCQPs having problem data that is both *acylcic* and *off-diagonally linearly separable from the origin*, are polynomial-time solvable.

Remark 4. *In [13], the authors impose the additional assumption that the primal feasible set is bounded to ensure finiteness of the optimal value. We remark that this assumption is not necessary, as Assumption 3.2.1 implies that both the primal and the dual problems have nonempty compact sets of optimal solutions [82].* ■

Given a QCQP with arbitrary graph structure, Theorem 3.2.4 points to a natural graph-structuring of a perturbation on the problem data to force a rank one optimal solution to the perturbed semidefinite relaxation of the QCQP. Namely, design the perturbation in order that the perturbed semidefinite program possesses acyclic graph structure and is off-diagonally linearly separable from the origin. We make this notion precise in our following definition of an *acyclic approximation* of a semidefinite program.

Definition 3.2.5 (Acyclic Approximation)**.** *The semidefinite program $\widetilde{d}$ is an (α, β)-acyclic approximation of the semidefinite program d if*

(i) the graph $G(\widetilde{d})$ is acyclic,

(ii) the data $\widetilde{d}$ is off-diagonally linearly separable from the origin,

(iii) $\|\widetilde{\mathcal{A}} - \mathcal{A}\| \leq \alpha$, $\|\widetilde{A}_0 - A_0\|_F \leq \beta$, and $\widetilde{b} = b$. ■

Definition 3.2.5 raises several interesting questions. For instance, when is it possible to construct an acyclic approximation of an arbitrary semidefinite program such that the *perturbed primal feasible set is a nonempty subset of the original primal feasible set*. Further, can one obtain *bounds on the performance* of primal optimal solutions obtained from the perturbed semidefinite program? We provide answers to these questions in our main results, Lemma 3.3.2 and Theorem 3.3.3.

3.3 Approximation Guarantess

We consider perturbations on a given semidefinite program $d = (\mathcal{A}, b, A_0)$ of the form:

$$\widetilde{\mathcal{A}} = \mathcal{A} + \Delta\mathcal{A} \quad \text{and} \quad \widetilde{A_0} = A_0 + \Delta A_0,$$

where $\Delta\mathcal{A} : \mathbf{H}^n \to \mathbf{R}^m$ is a linear map defined as

$$\Delta\mathcal{A}(X) := \Big[\mathrm{tr}(\Delta A_1 X), \ldots, \mathrm{tr}(\Delta A_m X)\Big]^\top$$

and $\Delta A_0, \Delta A_1, \ldots, \Delta A_m \in \mathbf{H}^n$. Denote the perturbed semidefinite program by $\widetilde{d} = (\widetilde{\mathcal{A}}, b, \widetilde{A}_0)$.

3.3.1 Distance to Infeasibility

When perturbing the problem data, it is essential to maintain the well posedness of the semidefinite program. In what follows, we review the notion of the *distance to infeasibility* of a conic linear system as defined by Renegar in [67].

Essentially, the distance to infeasibility (of the primal or dual problem) is the smallest perturbation in the data that yields an inconsistent system. To make this notion precise, we define $\mathcal{I}_P := \{d \mid \mathcal{P}(d) = \emptyset\}$ as the set of *primal infeasible* problem instances and $\mathcal{I}_D := \{d \mid \mathcal{D}(d) = \emptyset\}$ as the set of *dual infeasible* problem instances. We have the following definition from [67].

Definition 3.3.1 (Distance to Infeasibility)**.** *The distance from d to the set of primal infeasible instances is defined as*

$$\mathit{dist}(d, \mathcal{I}_P) := \inf\{\|d - \hat{d}\|_\pi \mid \hat{d} \in \mathcal{I}_P\},$$

and the distance from d to the set of dual infeasible instances is defined as

$$\mathit{dist}(d, \mathcal{I}_D) := \inf\{\|d - \hat{d}\|_\pi \mid \hat{d} \in \mathcal{I}_D\}.$$

Here, $\|\cdot\|_\pi$ denotes the product norm, defined as $\|d\|_\pi := \max\{\|\mathcal{A}\|, \|b\|_2, \|A_0\|_F\}$.

3.3.2 Characterizarion of Acyclic Approximations

We employ the notion of the distance to infeasibility to characterize a family of perturbations that yield nonempty *inner* acyclic approximations of a given semidefinite program d. More precisely, we provide sufficient conditions for the existence of an acyclic approximation $\widetilde{d}$ that satisfies $\emptyset \neq \mathcal{P}(\widetilde{d}) \subseteq \mathcal{P}(d)$. The sufficient conditions also guarantee that a feasible solution to the original QCQP can be recovered in polynomial time from an optimal solution to the perturbed semidefinite program $\widetilde{d}$. We also derive explicit bounds on the performance of said solution relative to the optimal value of the semidefinite relaxation. The following Lemma 3.3.2 and Theorem 3.3.3 constitute our main results. Suppose that Assumption 3.2.1 holds.

Lemma 3.3.2. *Consider a semidefinite program d having an (α, β)-acyclic approximation $\widetilde{d}$. If $\alpha <$ dist$(d, \mathcal{I}_P)$ and $\Delta A_0, \Delta A_1, \ldots, \Delta A_m \succeq 0$, then the following properties hold.*

(i) The perturbed semidefinite program $\widetilde{d}$ has a nonempty primal feasible set*, i.e.,*

$$\mathcal{P}(\widetilde{d}) \neq \emptyset.$$

(ii) The perturbed semidefinite program $\widetilde{d}$ is a primal inner *and a* dual outer approximation *of the original semidefinite program d, i.e.,*

$$\mathcal{P}(\widetilde{d}) \subseteq \mathcal{P}(d) \quad \text{and} \quad \mathcal{D}(\widetilde{d}) \supseteq \mathcal{D}(d).$$

(iii) A rank one primal optimal solution $\widetilde{X}$ to the perturbed semidefinite program $\widetilde{d}$ can be computed in polynomial time.

Proof. (i) Nonemptiness of $\mathcal{P}(\widetilde{d})$ is equivalent to dist$(\widetilde{d}, \mathcal{I}_P) > 0$, which we now show. Let $d = (\mathcal{A}, b, C)$ and define the projection $\Pi_P(d) := (\mathcal{A}, b, 0)$. It is obvious that

$$\text{dist}(d, \mathcal{I}_P) = \inf\{\|\Pi_P(d - \hat{d})\|_\pi \mid \hat{d} \in \mathcal{I}_P\}.$$

It follows from the previous fact and the reverse triangle inequality that

$$\text{dist}(\widetilde{d}, \mathcal{I}_P) \geq \text{dist}(d, \mathcal{I}_P) - \|\Pi_P(\widetilde{d} - d)\|_\pi.$$

Define $\Delta d := \widetilde{d} - d$. It follows that $\Delta d = (\Delta\mathcal{A}, 0, \Delta C)$. Hence, $\|\Pi_P(\widetilde{d} - d)\|_\pi = \|\Delta\mathcal{A}\|$. Finally, we have that

$$\text{dist}(\widetilde{d}, \mathcal{I}_P) \geq \text{dist}(d, \mathcal{I}_P) - \|\Delta\mathcal{A}\| \geq \text{dist}(d, \mathcal{I}_P) - \alpha > 0,$$

where the last two inequalities are a consequence of the Lemma assumptions.

(ii) We first prove that $\mathcal{P}(\widetilde{d}) \subseteq \mathcal{P}(d)$. Let $X \in \mathcal{P}(\widetilde{d})$ be arbitrary. It suffices to show that $\mathcal{A}(X) \le b$. Consider the following string of arguments:

$$\mathcal{A}(X) \;=\; (\widetilde{\mathcal{A}} - \Delta\mathcal{A})(X) \;\le\; b - \Delta\mathcal{A}(X) \;\le\; b.$$

The first inequality follows from the fact that $X \in \mathcal{P}(\widetilde{d})$. The second inequality follows from the assumption that $\Delta A_k \succeq 0$, as this implies that $\Delta\mathcal{A}(X) \ge 0$ for all $X \succeq 0$.

We now prove that $\mathcal{D}(\widetilde{d}) \supseteq \mathcal{D}(d)$. Let $y \in \mathcal{D}(d)$ be arbitrary. It suffices to show that $\widetilde{A}_0 + \widetilde{\mathcal{A}}^*(y) \succeq 0$. Consider the following string of arguments.

$$\widetilde{A}_0 + \widetilde{\mathcal{A}}^*(y) = A_0 + \mathcal{A}^*(y) + \Delta A_0 + \Delta\mathcal{A}^*(y) \succeq \Delta A_0 + \Delta\mathcal{A}^*(y) \succeq 0.$$

The first inequality follows from the assumption that $y \in \mathcal{D}(d)$. And the second inequality is an immediate consequence of $y \ge 0$ and $\Delta A_0, \Delta A_1, \ldots \Delta A_m \succeq 0$. This completes the proof of part (ii).

(iii) Consider the perturbed semidefinite program $\widetilde{d}$. By assumption, the data $\widetilde{d}$ is off-diagonally linearly separable from the origin and $G(\widetilde{d})$ is acyclic. Thus, by Theorem 3.2.4, it suffices to show that the semidefinite program $\widetilde{d}$ satisfies Assumption 3.2.1. Assumption 3.2.1 requires that $\mathcal{D}(\widetilde{d})$ is strictly feasible and $\mathcal{P}(\widetilde{d})$ is nonempty. We previously showed that $\mathcal{P}(\widetilde{d})$ is nonempty. To see why $\mathcal{D}(\widetilde{d})$ is strictly feasible, recall that $\mathcal{D}(\widetilde{d}) \supseteq \mathcal{D}(d)$. The result follows, as $\mathcal{D}(d)$ is strictly feasible by Assumption 3.2.1. ■

Essentially, Lemma 3.3.2 provides conditions under which a rank one primal feasible solution to the original semidefinite program d can be computed in polynomial time. Theorem 3.3.3 provides a data dependent performance guarantee.

Theorem 3.3.3. *Suppose the conditions of Lemma 3.3.2 hold. Let OPT denote the optimal value of the semidefinite program d. Then every primal optimal solution $\widetilde{X}$ to the perturbed* *semidefinite program $\widetilde{d}$ satisfies*

$$\mathit{OPT} \leq \operatorname{tr}(A_0 \widetilde{X}) \leq \mathit{OPT} + \varphi(d, \widetilde{d}), \tag{3.4}$$

where the error function $\varphi(d, \widetilde{d})$ is given by

$$\varphi(d, \widetilde{d}) := \max\{\|b\|_2, \mathit{OPT}\} \left(\mu + \nu \max\{\|A_0\|_F + \beta, -\mathit{OPT}\}\right),$$

where

$$\mu := \frac{\beta}{\mathit{dist}(d, \mathcal{I}_D)}, \quad \textit{and} \quad \nu := \frac{\alpha}{(\mathit{dist}(d, \mathcal{I}_P) - \alpha)\mathit{dist}(d, \mathcal{I}_D)}.$$

Proof. We first establish that OPT $\leq \operatorname{tr}(\widetilde{A}_0 \widetilde{X})$. To see why this is true, note that

$$\operatorname{tr}(\widetilde{A}_0 \widetilde{X}) = \operatorname{tr}(A_0 \widetilde{X}) + \operatorname{tr}(\Delta A_0 \widetilde{X}) \geq \operatorname{tr}(A_0 \widetilde{X}) \geq \mathrm{OPT}, \tag{3.5}$$

where the first inequality follows from the fact that $\Delta A_0 \succeq 0$ and the second inequality follows from the fact that $\widetilde{X}$ is feasible for the semidefinite program d.

The upper bound follows largely from arguments in Proposition 3.10 of [67]. In the proof, we have to distinguish between the optimal value of the primal and the dual problems. Let $J_P(d)$ denote the optimal value of the primal problem and $J_D(d)$ the optimal value of the dual problem of semidefinite program d.

Let $y \in \mathcal{D}(\widetilde{d})$ be a dual optimal solution to the semidefinite program $\widetilde{d}$. And consider another perturbation $d' := (\mathcal{A}, b, A_0')$, where

$$A_0' := A_0 + \Delta A_0',$$

and $\Delta A_0' := \Delta A_0 + \Delta \mathcal{A}^*(y)$. By nature of its construction, we have that $y \in \mathcal{D}(d')$, which implies that $-b^\top y \leq J_D(d')$. It follows readily that

$$J_D(\widetilde{d}) - J_D(d') \;\leq\; -b^\top y + b^\top y \;=\; 0.$$

And, as an immediate consequence, we have an intermediate bound on the difference between the primal optimal values of the original and perturbed problems:

$$J_D(\widetilde{d}) - J_D(d) \leq J_D(d') - J_D(d). \tag{3.6}$$

Let $X \in \mathcal{P}(d)$ be a primal optimal solution of the semidefinite program d. Since X is feasible for the semidefinite program d', it follows that

$$J_P(d') \leq \text{tr}(A_0'X) = J_P(d) + \text{tr}(\Delta A_0'X).$$

Hence, $J_P(d') - J_P(d) \leq \text{tr}(\Delta A_0'X)$. Since strong duality holds for the semidefinite program d, we must have $J_P(d) = J_D(d)$. Moreover, weak duality implies that $J_D(d') \leq J_P(d')$. It follows that

$$J_D(d') - J_D(d) \;\leq\; \text{tr}(\Delta A_0'X) \;\leq\; \|\Delta A_0'\|_F \|X\|_F, \tag{3.7}$$

where the second inequality follows from the Cauchy-Schwarz inequality. Furthermore, by using norm subadditivity, the definition of the operator norm, and the fact that $\|\Delta\mathcal{A}^*\| = \|\Delta\mathcal{A}\|$, we have that

$$\begin{aligned} \|\Delta A_0'\|_F &\leq \|\Delta A_0\|_F + \|\Delta\mathcal{A}^*(y)\|_F \\ &\leq \|\Delta A_0\|_F + \|\Delta\mathcal{A}\|\|y\|_2 \\ &\leq \beta + \alpha\|y\|_2. \end{aligned} \tag{3.8}$$

By combining (3.6), (3.7), and (3.8), we obtain

$$J_D(\widetilde{d}) - J_D(d) \leq \beta\|X\|_F + \alpha\|y\|_2\|X\|_F. \tag{3.9}$$

Since strong duality holds for the semidefinite program d, we must have $J_D(d) = J_P(d) =$ OPT. Similarly, since strong duality holds for semidefinite program $\widetilde{d}$, it must be true that $J_D(\widetilde{d}) = J_P(\widetilde{d}) = \text{tr}(\widetilde{A}_0\widetilde{X})$. By substituting these relations back to (3.9), we have

$$\text{tr}(\widetilde{A}_0\widetilde{X}) - \text{OPT} \leq \beta\|X\|_F + \alpha\|y\|_2\|X\|_F. \tag{3.10}$$

We will now bound $\|X\|_F$ and $\|y\|_2$. In [67], Renegar shows that if $\bar{X} \in \mathcal{P}(d)$ is primal optimal for semidefinite program d and $\bar{y} \in \mathcal{D}(d)$ is dual optimal for semidefinite program d, then

$$\|\bar{X}\|_F \;\leq\; \frac{\max\{\|b\|_2, J_P(d)\}}{\text{dist}(d, \mathcal{I}_D)}, \qquad \|\bar{y}\|_2 \;\leq\; \frac{\max\{\|A_0\|_F, -J_D(d)\}}{\text{dist}(d, \mathcal{I}_P)}. \tag{3.11}$$

Since $X \in \mathcal{P}(d)$ is an optimal solution for semidefinite program d and $y \in \mathcal{D}(\widetilde{d})$ is an optimal solution for semidefinite program $\widetilde{d}$, it follows from (3.10), (3.11), and inequality $\text{tr}(\widetilde{A}_0\widetilde{X}) \geq \text{OPT}$ (cf. Equation (3.5)) that

$$\text{tr}(\widetilde{A}_0\widetilde{X}) - \text{OPT} \;\leq\; \frac{\beta \max\{\|b\|_2, \text{OPT}\}}{\text{dist}(d, \mathcal{I}_D)} + \frac{\alpha \max\{\|b\|_2, \text{OPT}\} \max\{\|A_0 + \Delta A_0\|_F, -\text{OPT}\}}{\text{dist}(d, \mathcal{I}_D)\text{dist}(\widetilde{d}, \mathcal{I}_P)}.$$

To finish the proof, note that $\|A_0 + \Delta A_0\|_F \leq \|A_0\|_F + \|\Delta A_0\|_F \leq \|A_0\|_F + \beta$, where the first inequality follows from subadditivity and the second inequality follows from the assumption that the semidefinite program $\widetilde{d}$ is an (α, β)-acyclic approximation of the semidefinite program d. Moreover,

$$\text{dist}(\widetilde{d}, \mathcal{I}_P) \geq \text{dist}(d, \mathcal{I}_P) - \|\Delta \mathcal{A}\| \geq \text{dist}(d, \mathcal{I}_P) - \alpha,$$

where the above string of inequalities follows from the proof of part (i) of Lemma 3.3.2. Therefore,

$$\text{tr}(\widetilde{A}_0\widetilde{X}) \;\leq\; \text{OPT} + \mu \max\{\|b\|_2, \text{OPT}\} + \nu \max\{\|b\|_2, \text{OPT}\} \max\{\|A_0\|_F + \beta, -\text{OPT}\},$$

where μ and ν are as defined in the Theorem statement. To complete the proof note that by (3.5), $\text{tr}(A_0\widetilde{X}) \leq \text{tr}(\widetilde{A}_0\widetilde{X})$. ∎

The structure of the performance guarantee in (3.4) offers some interesting insight. First, notice that the error function is monotone in (α, β) and converges to zero as (α, β) tend to zero. This behavior is reassuring, as it implies good approximation guarantees for QCQPs with problem data that is *nearly acyclic*. Moreover, we recover the results of Theorem 3.2.4 for $(\alpha, \beta) = (0, 0)$. Second, the better conditioned the original semidefinite program – namley, the larger its primal distance to infeasibility – the better the performance guarantee. These observations are made transparent if we consider a semidefinite program d whose optimal value satisfies OPT $\geq \|b\|_2$. Under such assumption, the performance guarantee (3.4) simplifies to

$$\text{OPT} \leq \text{tr}(A_0 \widetilde{X}) \leq \text{OPT} \left(1 + \frac{\alpha \|A_0\|_F + \beta \text{dist}(d, \mathcal{I}_P)}{(\text{dist}(d, \mathcal{I}_P) - \alpha)\text{dist}(d, \mathcal{I}_D)}\right).$$

Finally, it bears mentioning that the performance guarantee (3.4) is data dependent and, therefore, does not provide a uniform bound for all QCQPs with general indefinite quadratic constraints.

Approximating the Distance to Infeasibility

The results of both Lemma 3.3.2 and Theorem 3.3.3 depend explicitly on the primal and dual distances to infeasibility of the original semidefinite program d. Their exact calculation is, in general, not tractable. They can be approximated, however. Using a conic theorem of alternatives, Freund et al. [23], formulate a tractable convex optimization problem whose optimal value approximates the primal distance to infeasibility. We have the following result from [23].

Proposition 3.3.4. Let $\mathbf{1} \coloneqq [1, \dots, 1]^\top \in \mathbf{R}^m$ and define $w \coloneqq \frac{1}{\sqrt{m}}\mathbf{1}$. The optimal value, $v(d)$, of the semidefinite program

$$\begin{aligned} v(d) \coloneqq \underset{\substack{g \in \mathbf{R},\ z \in \mathbf{R}^m \\ Q \in \mathbf{H}^n}}{\text{minimum}} \quad & \max\left\{\|\mathcal{A}^*(z) - Q\|_F,\ |b^\top z + g|\right\} \\ \text{subject to} \quad & w^\top z = 1, \\ & Q \succeq 0, \quad z \geq 0, \quad g \geq 0, \end{aligned}$$

approximates $\text{dist}(d, \mathcal{I}_P)$. Namely,

$$\frac{1}{\sqrt{m}} v(d) \;\leq\; \text{dist}(d, \mathcal{I}_P) \;\leq\; v(d).$$

■

One can follow a similar line of reasoning to derive analogous upper and lower bounds on the dual distance to infeasibility of the semidefinite program d. We have the following result.

Proposition 3.3.5. Let $I \in \mathbf{H}^n$ be the identity matrix and define $W := \frac{1}{\sqrt{n}} I$. The optimal value, $u(d)$, of the semidefinite program

$$\begin{aligned} u(d) := \underset{\substack{g \in \mathbf{R},\; q \in \mathbf{R}^m \\ Z \in \mathbf{H}^n}}{\text{minimum}} \quad & \max\left\{\|\mathcal{A}(Z) + q\|_2,\; |\text{tr}(A_0 Z) + g|\right\} \\ \text{subject to} \quad & \text{tr}(WZ) = 1, \\ & q \geq 0, \quad Z \succeq 0, \quad g \geq 0, \end{aligned}$$

approximates $\text{dist}(d, \mathcal{I}_D)$. Namely,

$$\frac{1}{\sqrt{n}} u(d) \;\leq\; \text{dist}(d, \mathcal{I}_D) \;\leq\; u(d).$$

■

Propositions 3.3.4 and 3.3.5 enable the reformulation of the conditions and results of Lemma 3.3.2 and Theorem 3.3.3 in a *computationally tractable*, albeit more conservative, form. We leave the details to the reader.

3.4 Conclusions

We have considered the design of a semidefinite programming-based approximation algorithm for a class of nonconvex quadratically constrained quadratic programs (QCQPs).

Recent work has shown that semidefinite relaxations of QCQPs with problem that is both acyclic and off-diagonally linearly separable yield optimal solutions that can be efficiently mapped to an optimal solution of the nonconvex QCQP . For general QCQPs, however, no such guarantees exist. In this chapter, we investigate the extent to which *structured perturbations* to the problem data can yield a perturbed QCQP whose feasible set is a *nonempty* subset of the original feasible set satisfying certain technical conditions and whose collective sparsity pattern defines an *acyclic graph*. Leveraging on the notion of the *distance to infeasibility* of a conic linear system, we provide a sufficient condition under which a perturbation is guaranteed to yield a nonempty feasible subset of the original feasible set. Moreover, we provide a computable bound on the degree of suboptimality incurred by a solution to the perturbed problem. The approximation guarantee depends on the problem data and the size of the perturbation.

CHAPTER 4

RECURSIVE SEMIDEFINITE APPROXIMATIONS OF AC-OPF

4.1 Introduction

In Chapter 2, we saw that the AC-OPF problem can be reformulated as a rank one constrained semidefinite program of the following form:

$$
\begin{aligned}
\underset{V \in \mathbf{H}^n}{\text{minimize}} \quad & \text{tr}(A_0 V) \\
\text{subject to} \quad & \text{tr}(A_i V) \le b_i, \qquad i = 1, \dots, m, \\
& \text{tr}(A_i V) = b_i, \qquad i = m+1, \dots, m+\ell, \\
& V \succeq 0, \\
& \text{rank}(V) \le 1,
\end{aligned}
\tag{4.1}
$$

The scalars $b_1, \dots, b_{m+\ell}$ and the Hermitian matrices $A_0, A_1, \dots, A_{m+\ell}$ are the given problem data, which capture power balance constraints and operational constraints reflecting bounds on real and reactive power generation, branch flows, and voltage magnitudes. The variable V is a Hermitian matrix, that is taken to be equal to the outer product of the vector of bus voltage phasors with itself. The fundamental difficulty in problem (4.1) is concentrated in the rank one constraint, which constitutes the only source of nonconvexity. The semidefinite relaxation of problem (4.1) entails removing the rank constraint.

Most commercial solvers implementing semidefinite programs (SDPs) rely on primal-dual interior point methods. Interior point methods for semidefinite programs are guaranteed to converge to a primal-dual optimal solution pair of maximal rank [50, 27]. More precisely, let $\mathcal{F}$ denote the feasible set for a given SDP. And, denote by $\mathcal{F}^\star \subseteq \mathcal{F}$ and $\text{relint}(\mathcal{F}^\star)$

the primal optimal face and its relative interior, respectively. The following result from [27] establishes that points belonging to the relative interior of the optimal face have maximum rank among all optimal solutions of the semidefinite program and that interior point methods are guaranteed to converge to optimal solutions in the relative interior of the optimal face.

Theorem 4.1.1. *[27, Lemma 3.1, 4.2] For any* $V \in \mathcal{F}^\star$ *and* $W \in \operatorname{relint}(\mathcal{F}^\star)$*,* $\operatorname{col}(V) \subseteq \operatorname{col}(W)$*. In other words,*

$$\operatorname{rank}(W) = \max\{\operatorname{rank}(V) : V \in \mathcal{F}^\star\}, \quad \textit{for all } W \in \operatorname{relint}(\mathcal{F}^\star).$$

Moreover, interior point methods for semidefinite programs converge to an optimal solution $W \in \operatorname{relint}(\mathcal{F}^\star)$.

The implication of Theorem 4.1.1 is that the semidefinite relaxation of the AC-OPF problem will fail to yield an optimal solution that can be efficiently mapped back to the original feasible solution set if the optimal face of the semidefinite relaxation contains points with rank strictly greater than one. With the aim of quantifying the role of optimal facial structure in either realizing or obfuscating efficiency of the semidefinite relaxation, we delineate the following three categories of optimal facial geometries.

C1. The maximal rank of the optimal face is one.

C2. The minimal rank of the optimal face is strictly greater than one.

C3. The minimal rank of the optimal face is one, while the maximal rank is strictly greater than one.

Category C1 will have *only rank one optimal solutions,*

$$\max\{\operatorname{rank}(V) : V \in \mathcal{F}^\star\} = 1.$$

In this case, the semidefinite relaxation will yield a rank one optimal matrix that can be mapped, through a dyadic decomposition, to a globally optimal solution of the original AC-OPF problem.

Category C2 corresponds to semidefinite relaxations that do not admit rank one optimal points. Namely, the minimal rank of the optimal face is strictly greater than one

$$\min\{\operatorname{rank}(V) : V \in \mathcal{F}^\star\} > 1,$$

which implies that the optimal value of such a semidefinite relaxation would yield a strict lower bound on the global minimum of the original AC-OPF problem (2.6). This amounts to a nonzero optimality gap between the relaxation and the original problem. Clearly then, verifying strictness of the global lower bound given by the semidefinite relaxation amounts to verifying emptiness of the intersection between the optimal face and the set of all rank one positive semidefinite matrices.

This condition has a natural geometric interpretation for matrices belonging to the positive semidefinite cone $\mathbf{H}_+^n$. Namely, a matrix $V \in \mathbf{H}_+^n$ has rank one if and only if it spans an *extreme ray* of the cone [5]. Hence, the semidefinite relaxation will possess a rank one optimal solution if and only if its optimal face $\mathcal{F}^\star$ has a nonempty intersection with and extreme ray of $\mathbf{H}_+^n$. In Section 2.5.2, we developed an a posteriori sufficient condition for checking that the semidefinite relaxation of AC-OPF is inexact by verifying uniqueness of its optimal solution. A different way for verifying inexactness is provided by the Positivstellensatz.

Remark 5. (Positivstellensatz). *The nonexistence of rank one optimal solutions to the* *semidefinite relaxation can be verified numerically by means of a Positivstellensatz-based infeasibility certificate. Stengle's Positivstellensatz states that if a system of polynomial equations and inequalities defining a semialgebraic set is infeasible, it is always possible to find an algebraic certificate that confirms that the said semialgebraic set is empty [11]. The construction of polynomials that satisfy said identity can be accomplished through sum of squares programming with bounded degree polynomials on the semialgebraic set defined by the intersection of the optimal face of the semidefinite relaxation with the rank one algebraic variety. One drawback of this approach, however, is that the computational complexity required to implement such sum of squares methods grows rapidly as a function of the number of constraints, variables, and degree of polynomials.*

Remark 6. (A 3-bus system with no rank one solutions). *There exist exceedingly simple power systems whose semidefinite relaxations do not admit a rank one optimal solution. Consider, for example, the three bus system examined in [43]. One can readily verify, through an application of Theorem 2.5.8 that the optimal value of the semidefinite relaxation is a strict lower bound on the global optimum of the AC-OPF problem. Moreover, this example gives pause, as it reveals the potential fragility of such relaxations. Further theoretical work is required to provide general a priori sufficient conditions under which the semidefinite relaxation of AC-OPF is guaranteed to fail.*

Category C3 refers to the family of semidefinite relaxations possessing both high rank and *hidden* rank one optimal solutions. More precisely,

$$\min\{\text{rank}(X) : X \in \mathcal{F}^\star\} = 1 \quad \text{and}$$
$$\max\{\text{rank}(X) : X \in \mathcal{F}^\star\} > 1.$$

We refer to the rank one solutions as *hidden*, given the propensity of interior point methods to converge to optimal points of maximal rank (cf. Theorem 4.1.1). A solution to a semidefinite relaxation belonging to this family will fail to yield useful information regarding the potential optimality gap induced by the relaxation. In certain cases, the solution to the semidefinite relaxation can be efficiently mapped back to feasible set of the AC-OPF problem without loss of optimality (cf. Section 2.5.1). Recent results, for example, have shown that AC-OPF problems, satisfying certain technical conditions and defined on networks with acyclic topologies, yield semidefinite relaxations with at most rank one optimal solutions (cf. Section 2.5.1 and [13, 94]). In general, however, mapping a high-rank solution to the semidefinite relaxation back to the original feasible set is NP-hard.

This inspires the exploration of methodologies capable of uncovering *hidden* rank one optimal solutions to the semidefinite relaxation, when they exist. Qualitatively, this amounts to identifying matrices of minimal rank among all matrices belonging to the optimal face of the semidefinite program (2.10). The optimal face is defined as

$$\mathcal{F} = \{V \in \mathcal{F} \ : \ \text{tr}(A_0 V) = J^\star\}, \tag{4.2}$$

where $J^\star$ denotes the optimal value of the semidefinite relaxation. Essentially, computing an optimal matrix of minimal rank entails the solution of a rank minimization problem restricted to the optimal face of the semidefinite relaxation.

$$\begin{aligned} \underset{V \in \mathbf{H}^n}{\text{minimize}} \quad & \text{rank}(V) \\ \text{subject to} \quad & V \in \mathcal{F}^\star \end{aligned} \tag{4.3}$$

Remark 7 (AC-OPF as Rank Minimization)**.** *In the event that the optimal face of the semidefinite relaxation possesses a rank one matrix, problem (4.3) reveals that AC-OPF can be equivalently reformulated as rank minimization problem over a spectrahedral set. Explicit rank minimization, however, is known to be NP-hard in general.*

As a tractable alternative, one might naturally solve an approximation to the rank minimization problem through suitable choice of a convex surrogate for rank, which is neither continuous nor convex.

In [21], Fazel et al. prove that the nuclear norm is the convex envelope of rank on spectral norm balls. This property fails to hold, however, for general convex sets. While the nuclear norm has been shown to be an effective surrogate for rank over certain affine equality constrained sets satisfying a restricted isometry property [66], it can behave quite poorly over more general spectrahedral sets. In fact, when optimizing over the feasible spectrahedron derived from the semidefinite relaxation of the AC-OPF problem, one can show that naive nuclear norm minimization will frequently fail to find *low-rank* feasible solutions – even when they exist. This behavior derives from the near invariance of nuclear norm over the feasible set – an observation also made by the authors in [52]. More precisely, for any feasible matrix V, one can readily derive the following lower and upper bounds on its nuclear norm

$$\sum_{i=1}^{n}(v_i^{\min})^2 \;\leq\; \|V\|_* \;\leq\; \sum_{i=1}^{n}(v_i^{\max})^2, \tag{4.4}$$

where, $\|V\|_*$ denotes the nuclear norm of V (see Section 4.2 for the definition of nuclear norm). In practice, the lower and upper bounds on bus voltage magnitude – $v_i^{\min}$ and $v_i^{\max}$, respectively – are chosen to be close to 1 per unit for all buses i, because of strict requirements on power quality. This suggests that all feasible solutions to the semidefinite relaxation of AC-OPF (2.10) have nearly equal nuclear norm, which reveals why naive nuclear norm regularization may fail to distinguish between *low* and *high-rank* solutions.

Contribution: As an alternative to nuclear norm minimization, we analyze in Section 4.3 the behavior of an algorithm that involves solving a sequence of weighted trace minimization problems, where the weighting matrices are recursively chosen to drive small (but nonzero) eigenvalues of the successive solution iterates to zero – an approach which derives largely from the work in [22]. In addition, a simple bisection method is proposed in Section 4.4.2 to address problems for which the above procedure fails to yield a rank-one optimal solution. The algorithms are tested on multiple representative power system examples. In many cases, the weighted trace minimization heuristic obtains a hidden rank-one solution, where the naive semidefinite relaxation fails.

In addition to these two algorithms, we provide in Section 4.2 an equivalent representation for the rank one inequality constraint of problem (4.1) as the difference of two convex functions. Using this representation, we develop in Section 4.2.1 a convex inner approximation to the rank one constrained semidefinite program (4.1) and an algorithm, which is guaranteed to generate a sequence of feasible solutions that have nonincreasing costs. However, a feasible solution to the rank one constrained semidefinite program is required to initiate the algorithm. We provide a heuristic for a computing such a feasible point.

4.2 Equivalent Representation of Rank

In this Section, we develop an equivalent representation for the rank one inequality constraint in problem (4.1). This equivalent representation is described by a function which can be decomposed as the difference of two convex functions and it is obtained by representing the rank of a matrix in terms of its nonzero singular values. This equivalent representation of a rank one constraint lends itself to a convex inner approximation for the rank one constrained semidefinite program, which we present is Section 4.2.1.

For a matrix $V \in \mathbf{H}^n$, let $\sigma_i(V)$ be i^{th} largest singular value of V. In addition, let

$$\|V\|_* := \sum_{i=1}^{n} \sigma_i(V)$$

be the nuclear norm of V and

$$\|V\|_F := \sqrt{\operatorname{tr}(XX^\top)} = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}[V]_{ij}^2} = \sqrt{\sum_{i=1}^{n}\sigma_i(V)^2}, \tag{4.5}$$

be the Frobenius norm of V. If $V \in \mathbf{H}_+^n$, then the eigenvalues of X are equal to its singular values. Therefore the nuclear norm of a positive semidefinite matrix equals its trace. In addition, it is straightforward to verify that the Frobenius norm of a Hermitian positive semidefinite matrix X is no greater than its trace, i.e.,

$$\|V\|_F \leq \operatorname{tr}(V). \tag{4.6}$$

The rank of a matrix V is equal to the number of nonnegative singular values of V. More precisely,

$$\operatorname{rank}(V) = \sum_{i=1}^{n} \mathbf{1}_{\{\sigma_i(V)>0\}}, \tag{4.7}$$

where $\mathbf{1}_S$ is the indicator function of the set S. The above characterization of the rank of a matrix in terms of its singular values together with (4.6) give rise to the following equivalent representation of a rank one inequality constraint. We have the following Lemma.

Lemma 4.2.1. *Let $V \in \mathbf{H}_+^n$. Then,*

$$\operatorname{rank}(V) \leq 1 \quad \Longleftrightarrow \quad \operatorname{tr}(V) - \|V\|_F \leq 0.$$

Proof. First let $\operatorname{rank}(V) \leq 1$. We must show that $\operatorname{tr}(V) - \|V\|_F \leq 0$. Since $\operatorname{rank}(V) \leq 1$, it follows by the characterization of rank in equation (4.7) that

$$\sigma_2(V) = \sigma_3(V) = \cdots = \sigma_n(V) = 0.$$

Therefore tr$(V) = \sum_{i=1}^{n} \sigma_i(V) = \sigma_1(V)$ and $\|V\|_F = \sqrt{\sum_{i=1}^{n} \sigma_i(V)^2} = \sigma_1(V)$. This establishes the desired inequality.

Next, suppose that tr$(V) \leq \|V\|_F$. We must show that $\mathrm{rank}(V) \leq 1$. Since $V \in \mathbf{H}_+^n$, it follows that tr$(V) \geq 0$. Hence, we must also have tr$(V)^2 \leq \|V\|_F^2$. First observe that

$$\mathrm{tr}(V)^2 = \left(\sum_{i=1}^{n} \sigma_i(V)\right)^2 = \sum_{i=1}^{n} \sigma_i(V)^2 + 2\sum_{i=1}^{n} \sum_{j:j>i} \sigma_i(V)\sigma_j(V) = \|V\|_F^2 + 2\sum_{i=1}^{n} \sum_{j:j>i} \sigma_i(V)\sigma_j(V)$$

This implies that

$$\sum_{i=1}^{n} \sum_{j:j>i} \sigma_i(V)\sigma_j(V) \leq 0$$

And since $\sigma_i(V) \geq 0$, for all $i = 1, \ldots, n$ and $\sigma_1(V) \geq \sigma_2(V) \geq \cdots \geq \sigma_n(V)$, it is easy to verify that the above inequality implies that

$$\sigma_2(V) = \cdots = \sigma_n(V) = 0.$$

Therefore, we must have $\mathrm{rank}(V) \leq 1$. ■

The function tr$(V) - \|V\|_F$ in Lemma 4.2.1 is a concave function of V as it is the difference of a linear function of V and a convex function of V. In Section 4.2.1, we will use this equivalent characterization of the rank one constraint to develop an algorithm, which can be used to approximate the AC-OPF problem from within.

Lemma 4.2.1 gives rise to an equivalent reformulation for the rank one constrained semidefinite program (4.1) We have the following Corollary.

Corollary 4.2.2. *The optimization problem*

$$\begin{aligned}
\underset{V\in\mathbf{H}^n}{\text{minimize}} \quad & \text{tr}(A_0 V) \\
\text{subject to} \quad & \text{tr}(A_i V) \le b_i, \qquad i = 1,\dots,m, \\
& \text{tr}(A_i V) = b_i, \qquad i = m+1,\dots,m+\ell, \\
& V \succeq 0, \\
& \text{tr}(V)^2 - \|V\|_F^2 \le 0.
\end{aligned} \tag{4.8}$$

is equivalent to the rank one constrained semidefinite program (4.1).

In (4.8) the constraint $\text{tr}(V)^2 - \|V\|_F^2 \le 0$ is equivalent to $\text{tr}(V) - \|V\|_F \le 0$ since V is positive semidefinite (therefore, $\text{tr}(V) \ge 0$). This representation of the rank one constraint will be convenient in Section 4.2.1 when we develop an algorithm to approximate the feasible set of the AC-OPF problem from within.

4.2.1 Convex Inner Approximations

In this section, we leverage on the equivalent representation of the rank one inequality constraint described in Section 4.2 to develop an algorithm, which gives rise to a convex inner approximation for the rank one constrained semidefinite program (4.1). Our method generates a sequence of feasible solutions for (4.1) that have nonincreasing costs.

Recall the function $\text{tr}(V)^2 - \|V\|_F^2$ in Lemma 4.2.1, which is the difference of two convex functions. The crux of our approach is based on the simple observation that any concave function can be approximated from above with its linearization at a point. The sum of this linearization with the convex component of the original function yields a convex global overestimator of the original nonconvex function. More precisely, let $g : \mathbf{H}^n \to \mathbf{R}$ be a function given by $g(Z) := -\|Z\|_F^2$. Since g is concave, $g(Z)$ is smaller than its linearization around a point $W \in \mathbf{H}^n$. In particular,

$$-\|Z\|_F^2 = g(Z) \leq g(W) + \mathrm{tr}\big(\nabla g(W)^\top (Z - W)\big) = \|W\|_F^2 - 2\mathrm{tr}(WZ),$$

where $\nabla g : \mathbf{H}_+^n \to \mathbf{R}$ is the gradient of g. It follows that for all $W \in \mathbf{H}^n$

$$\mathrm{tr}(V)^2 - \|V\|_F^2 \leq \mathrm{tr}(V)^2 - 2\mathrm{tr}(WV) + \|W\|_F^2. \tag{4.9}$$

Let $f : \mathbf{H}^n \times \mathbf{H}^n \to \mathbf{R}$ be a function given by

$$f(V, W) := \mathrm{tr}(V)^2 - 2\mathrm{tr}\big(WV\big) + \|W\|_F^2.$$

Note that $f(V, W)$ is a convex function of V, for any fixed W and a convex function of W for any fixed V. Leveraging on the equivalent characterization of a rank one inequality constraint provided in Lemma 4.2.1, we have the following Corollary.

Corollary 4.2.3. *Let $V \in \mathbf{H}^n$ be a positive semidefinite matrix. Then, $\mathrm{rank}(V) \leq 1$ if and only if there exists a matrix $W \in \mathbf{H}^n$ such that*

$$f(V, W) \leq 0.$$

The above Corollary gives rise to a convex inner approximation for the rank one constrained semidefinite program (4.8).

Corollary 4.2.4. *Let $W \in \mathbf{H}^n$ and consider the following convex program*

$$\begin{array}{lll}
\underset{V \in \mathbf{H}^n}{\text{minimize}} & \mathrm{tr}(A_0 V) & \\
\text{subject to} & \mathrm{tr}(A_i V) \leq b_i, & i = 1, \ldots, m, \\
& \mathrm{tr}(A_i V) = b_i, & i = m+1, \ldots, m+\ell, \\
& V \succeq 0, & \\
& f(V, W) \leq 0. &
\end{array} \tag{4.10}$$

Algorithm 4.1: Alternate minimization Algorithm

Given a convex set $\mathcal{F}$, an initial condition $X^0 \in \mathbf{H}^n$, and maximum number of iterations $\overline{k}$.

Initialize $k = 0$

Repeat

1. *Compute.* $X^{k+1} \in \underset{X \in \mathcal{F}}{\text{argmin}}\ \text{tr}(X)^2 - 2\text{tr}\big(X^k X\big)$
2. *Update.* $k = k + 1$

Until $\mathrm{rank}(X^k) \leq 1$ or $k = \overline{k}$

Output X^{k+1}.

Table 4.1: Alternate minimization method for finding a positive semidefinite matrix $W \in \mathbf{H}^n$ which yields a nonempty feasible set for problem (4.10).

An optimal solution of (4.10) *is a feasible solution for the rank one constrained semidefinite program* (4.8).

Remark 8. *The constraint $f(V, W) \leq 0$ in problem* (4.10) *can be expressed as the following linear matrix inequality through a direct application of the Schur complement condition for positive semidefiniteness.*

$$\begin{bmatrix} 1 & \text{tr}(V) \\ \text{tr}(V) & 2\text{tr}(WV) - \|W\|_F^2 \end{bmatrix} \succeq 0.$$

For further details on the Schur complement, see Appendix B.6.

An important challenge to the implementation of the convexification technique developed above lies in computing a matrix $W \in \mathbf{H}^n$, which yields a nonempty feasible set for (4.10). We propose a heuristic for computing one such matrix. The corresponding algorithm is described in Table 4.1 and it amounts to an alternate minimization approach. In particular, the alternate minimization method aims to find a pair of matrices $(V, W) \in \mathcal{F} \times \mathbf{H}^n$ such that $f(V, W) \leq 0$. In essence, this amounts to computing a feasible solution to the rank

Algorithm 4.2: Linearization Minimization Algorithm

Given a matrix V^0

Initialize $k = 0$

Repeat

1. *Compute.* $V^{k+1} \in \underset{V\in\mathcal{F}(V^k)}{\text{argmin}}\ \text{tr}(A_0 V)$
2. *Update.* $k = k + 1$

Until $\|V^k - V^{k-1}\|_F \leq \varepsilon$.

Output $v = \sqrt{[\Lambda]_{11}}Ue_1$, where $V^k = U\Lambda U^*$ is an eigenvalue decomposition of V^k.

Table 4.2: Linearization-minimization algorithm, which yields a sequence of *feasible* solutions for the AC-OPF problem with nonincreasing costs.

one constrained semidefinite program (4.1). To do so, the algorithm minimizes $f(V, W)$ over $\mathcal{F}$, while keeping W fixed and then it minimizes $f(V, W)$ over $\mathbf{H}^n_+$, while keeping V fixed. More precisely, given an initial condition $W^0 \in \mathbf{H}^n$, we take

$$V^{k+1} \in \underset{V\in\mathcal{F}}{\text{argmin}}\ f(V, W^k) = \underset{V\in\mathcal{F}}{\text{argmin}}\ \text{tr}(V)^2 - 2\text{tr}\big(W^k V\big)$$

$$W^{k+1} = \underset{W\in\mathbf{H}^n}{\text{argmin}}\ f(V^{k+1}, W) = \underset{W\in\mathbf{H}^n}{\text{argmin}}\ \|W - V^{k+1}\|_F^2 = V^{k+1},$$

where the superscript k denotes the k^{th} iteration step. Algorithm 4.1 is not guaranteed to return a rank one feasible point for the rank constrained semidefinite program (4.1). If it does, however, we provide a linearization-minimization algorithm, which is guaranteed to yield a sequence of feasible point to the AC-OPF problem whose costs is nonincreasing. To describe the linearization-minimiaztion algorithm it will be convenient to define the feasible set of problem (4.8), which is parameterized by the matrix W as follows

$$\mathcal{F}(W) := \mathcal{F} \cap \{V \in \mathbf{H}^n \mid f(V, W) \leq 0\}.$$

We have the following Proposition, which establishes two important properties of the recursive method: (i) it is guaranteed to yield a nonempty convex inner approximation to the rank one constrained semidefinite program at each step in the recursion, provided that Algorithm 4.1 returns a feasible solution to the rank one constrained semidefinite program (4.1) and (ii) it is guaranteed to generate a sequence of *feasible* solutions for the AC-OPF problem with nonincreasing costs.

Proposition 4.2.5. *Let V^0 be a matrix returned by the alternating minimization Algorithm 4.1. In addition, let $\{V^k\}_{k=1}^{\infty}$ denote the sequence of solutions generated by the linearization-minimization Algorithm 4.2. The following properties hold for each step k of the recursion.*

(i) Nonemptiness: $\mathcal{F}(V^k) \neq \emptyset$.

(ii) Cost montonicity: $\text{tr}(A_0 V^k) \leq \text{tr}(A_0 V^{k-1})$.

Proof. The proof is by induction on k.

Base of induction: Consider the case $k = 0$ and suppose that $V^0 \in \mathbf{H}^n$ is a rank one matrix returned by the alternating minimization algorithm in Table 4.1. We claim that $V^0 \in \mathcal{F}(V^0)$. By construction of Algorithm 4.1, we have that $V^0 \in \mathcal{F}$. And since $\text{rank}(V^0) \leq 1$, we obtain through a direct application of Lemma 4.2.1 that

$$\text{tr}(V^0)^2 - 2\text{tr}(V^0 V^0) + \|V^0\|_F^2 = \text{tr}(V^0)^2 - \|V^0\|_F^2 \leq 0$$

Therefore, $V^0 \in \mathcal{F}(V^0)$ and we are done with (4.2.1) for the base of induction. Let

$$V^1 \in \underset{V \in \mathcal{F}(V^0)}{\text{argmin}} \; \text{tr}(A_0 V)$$

By its optimality, we must have $\text{tr}(A_0 V^1) \leq \text{tr}(A_0 V)$, for all $V \in \mathcal{F}(V^0)$. And since $V^0 \in \mathcal{F}(V^0)$, we obtain the desired inequality.

Step of Induction: Let $t = s$ and suppose that $\mathcal{F}(V^j) \neq \emptyset$ for all $j < s$. We must show that $\mathcal{F}(V^s) \neq \emptyset$ in order to establish (4.2.1). By the induction hypothesis, we have that

$\mathcal{F}(V^{s-1}) \neq \emptyset$. Let

$$V^s \in \underset{X \in \mathcal{F}(V^{s-1})}{\text{argmin}} \ \text{tr}(A_0 V).$$

We claim that $V^s \in \mathcal{F}(V^s)$. Since $X^s \in \mathcal{F}(V^{s-1})$, we have that $V^s \in \mathcal{F}$ and

$$\text{tr}(V^s)^2 - 2\text{tr}(V^{s-1}V^s) + \|V^{s-1}\|_F^2 \leq 0.$$

Therefore, by Lemma 4.2.1, we obtain that $\text{rank}(V^s) \leq 1$. Hence,

$$\text{tr}(V^2)^2 - 2\text{tr}(V^sV^s) + \|V^s\|_F^2 = \text{tr}(V^s)^2 - \|V\|_F^2 \leq 0.$$

It follows that $V^s \in \mathcal{F}(V^s)$, which establishes the induction step for (4.2.1). Since V^s is an optimal solution, we must have

$$\text{tr}(A_0V^s) \leq \text{tr}(A_0V),$$

for all $V \in \mathcal{F}(V^{s-1})$. Moreover, we established in part of the induction step that $V^k \in \mathcal{F}(V^k)$, for all $k = 0, 1, \dots$. Therefore,

$$\text{tr}(A_0V^s) \leq \text{tr}(A_0X^{s-1}).$$

This completes the proof of the induction step for part ((ii)). ■

4.3 Concave Approximations of Rank

In this Section, we consider the rank minimization problem (4.3) and we take the approach of approximating rank with a continuously differentiable, strictly concave function $g : \mathbf{H}_+^n \to \mathbf{R}$. With g acting as a surrogate for rank, we instead propose to solve the alternative problem

$$\begin{aligned} \text{minimize} \quad & g(V) \\ \text{subject to} \quad & V \in \mathcal{C} \end{aligned} \tag{4.11}$$

where $\mathcal{C} \subset \mathbf{H}_+^n$ is a convex, compact subset of the positive semidefinite cone. To address the nonconvexity of problem (4.11), in Section 4.3.1 we describe a standard iterative linearization-minimization algorithm to obtain a sequence of convex differentiable problems, whose optimal solutions are guaranteed to converge to a local minimum of g on $\mathcal{C}$. In Section 4.3.2, we focus our attention on specific instances of g belonging to the log-det family. Namely, we consider

$$g(V) = \log\det(f(V) + \delta I),$$

where the underlying parameterization (in the regularization constant $\delta > 0$ and mapping $f : \mathbf{H}_+^n \to \mathbf{H}_+^n$) controls the quality of g's approximation to rank. Notice, that for $f(V) = V$, we recover the classical log-det heuristic [22, 55]. Working with rank surrogates of this form, we employ a gradient descent method to compute a local minimum of g – with the aim of recovering a rank one matrix belonging to the optimal face $\mathcal{F}^\star$ of the semidefinite relaxation for AC-OPF (cf. Eq. (4.2)). In the event that we fail to recover a rank one matrix in the optimal face, we suggest in Section 4.3.3 a simple bisection algorithm to iteratively relax the set of feasible points until a rank one feasible point is obtained. Finally, in Section 4.3.4 we explore how one might iteratively choose a sequence of regularization parameters $\{\delta^k\}$, so that the resulting solution iterates satisfy certain *rank monotonicity* properties.

4.3.1 Iterative Linearization-Minimization Algorithm

In this section, we introduce the iterative linearization–minimization algorithm and we discuss its convergence properties. We work in a general framework where we consider arbitrary strictly concave functions and arbitrary convex compact subsets of the positive semidefinite cone.

Let $\{g^k\}_{k \in \mathbf{N}}$ be a sequence of smooth, strictly concave functions, converging pointwise to g over a convex, compact set $\mathcal{C} \subset \mathbf{H}_+^n$. Moreover, assume that the sequence is monotonic nonincreasing. Namely,

$$g^{k+1}(V) \leq g^k(V) \quad \text{for all } \ V \in \mathcal{C} \ \text{ and } \ k \in \mathbf{N}.$$

For each $k \in \mathbf{N}$, we define the linearization of $g^k(V)$ around $W \in \mathcal{C}$ as

$$\Lambda^k(V, W) := g^k(W) + \mathrm{tr}[\nabla g^k(W)^*(V - W)],$$

from which we readily derive the iterative linearization-minimization algorithm as follows.

$$V^{k+1} \in \underset{V \in \mathcal{C}}{\mathrm{argmin}} \ \Lambda^{k+1}(V, V^k) = \underset{V \in \mathcal{C}}{\mathrm{argmin}} \ \mathrm{tr}\left(\nabla g^{k+1}(V^k)^* V\right), \tag{4.12}$$

where $\nabla g : \mathbf{H}_+^n \to \mathbf{H}^n$ is the gradient of g. The algorithm can be initialized at any point $V^0 \in \mathbf{H}_+^n$. Before presenting the result on convergence, we have the following useful Lemma.

Lemma 4.3.1. *Consider a sequence of iterates $\{V^k\}$ generated by the recurrence relation (4.12). We have that*

$$g^{k+1}(V^{k+1}) < g^k(V^k), \tag{4.13}$$

for all $k \in \mathbf{N}$ such that $V^k \neq V^{k+1}$. And

$$\lim_{k \to \infty} \mathrm{tr}\left(\nabla g^{k+1}(V^k)^*(V^{k+1} - V^k)\right) = 0. \tag{4.14}$$

Proof. First consider the proof of (4.13). By strict concavity, we have that $g^k(V)$ is strictly less than it's linearization around $W \in \mathcal{C}$ for all $V \neq W$. In particular, for $V^k \neq V^{k+1}$, we have that

$$g^{k+1}(V^{k+1}) < g^{k+1}(V^k) + \mathrm{tr}(\nabla g^{k+1}(V^k)^*(V^{k+1} - V^k)).$$

And by optimality of V^{k+1} according to (4.12), we arrive at $g^{k+1}(V^{k+1}) < g^{k+1}(V^k)$. The desired result follows immediately from monotonicity of the sequence $\{g^k\}$.

Consider now the proof of (4.14). The sequence $\{g^k(V^k)\}$ of real numbers is bounded from below by continuity of the limit function g over a compact set $\mathcal{C}$. Hence, it follows from (4.13) and the monotone convergence theorem that the sequence $\{g^k(V^k)\}$ has a finite limit. The desired result follows from the fact that the quantity $g^{k+1}(V^k) + \mathrm{tr}(\nabla g^{k+1}(V^k)^\top(V^{k+1}-V^k))$ is sandwiched from above and below by $g^k(V^k)$ and $g^{k+1}(V^{k+1})$, respectively. ∎

We now discuss the convergence properties of the proposed algorithm. First, we provide the definition of a stationary point.

Definition 4.3.2. *Let $h : \mathbf{C}^{n\times n} \to \mathbf{R}$ be a continuously differentiable function defined on the set $\mathcal{S} \subseteq \mathbf{R}^{n\times n}$. A matrix $V \in \mathcal{S}$ satisfying*

$$\mathrm{tr}\left(\nabla h(V)^*(Y - V)\right) \geq 0 \quad \forall \;\; Y \in \mathcal{S}$$

is a said to be a stationary point *of h over $\mathcal{S}$.*

Theorem 4.3.3. *Consider a sequence of iterates $\{V^k\}$ generated by the recurrence relation (4.12). We have the following convergence properties.*

(i) The sequence $\{V^k\}$ satisfies $\|V^{k+1} - V^k\|_F \to 0$.

(ii) Every limit point of $\{V^k\}$ is a stationary point.

Proof. While the proof of Theorem 4.3.3 (i) follows largely from the proof of Theorem 14.1.3 in [62], we include a concise version here for completeness. First, we define a hemivariate functional and a strongly downward sequence. We then use these definitions to show that a strictly concave function is hemivariate and that the sequence of iterates $\{V^k\}$ generated as in (4.12) is strongly downward in the function g^k for each k.

Definition 4.3.4. [62,] A functional $g : \mathbf{R}^{n\times n} \to \mathbf{R}$ is said to be *hemivariate* on a set $\mathcal{S}_0 \subset \mathbf{R}^{n\times n}$ if it is not constant on any line segment of $\mathcal{S}_0$ – that is, if there does not exist distinct points $V, W \in \mathcal{S}_0$ such that $\theta V + (1-\theta)W \in \mathcal{S}_0$ and $g(\theta V + (1-\theta)W) = g(V)$ for all $\theta \in [0,1]$.

Definition 4.3.5. [62,] Let $g : \mathbf{C}^{n\times n} \to \mathbf{R}$ and $\{V^k\}$ be a sequence of iterates in some subset $\mathcal{S}_0 \subset \mathbf{C}^{n\times n}$. We say that $\{V^k\}$ is *strongly downwards* in g if:

(a) $\theta V^k + (1-\theta)V^{k+1} \in \mathcal{S}_0$

(b) $g(V^k) \;\geq\; g(\theta V^k + (1-\theta)V^{k+1}) \;\geq\; g(V^{k+1})$

for all $\theta \in [0,1]$.

We are now ready to show that the sequence of iterates $\{V^k\}$ generated by the iterative linearization-minimization (4.12) is strongly downward in g^k (for each k) and that the limit function g of the sequence of the $\{g^k\}$ is hemivariate.

Lemma 4.3.6. Let $\{g^k\}$ be a sequence of smooth, strictly concave functions converging pointwise to a smooth, strictly concave function g over a convex, compact set $\mathcal{C} \subset \mathbf{H}^n_+$. The following statements hold.

1. The limit function g is hemivariate.
2. A sequence of iterates $\{V^k\}$ generated by the iterative linearization-minimization

(4.12) is strongly downward in the function g^k for each k.

Proof:

1) Suppose, for the sake of contradiction, that $\exists$ distinct $V, W \in \mathcal{C}$ such that

$$g(\theta V + (1-\theta)W) = g(V) \quad \forall\, \theta \in [0,1]$$

Since g is strictly concave,

$$g(V) = g(\theta V + (1-\theta)W) > \theta g(V) + (1-\theta)g(W)$$

for all $\theta \in [0,1]$. By taking $\theta = 1$ we have $g(V) > g(V)$, a contradiction.

2) Let V^k, V^{k+1} be two successive iterates belonging to $\mathcal{C}$. Because $\mathcal{C}$ is convex, it follows that $V_a = \theta V^k + (1-\theta)V^{k+1} \in \mathcal{C}$ for all $\theta \in [0,1]$. By strict concavity, we have

$$\begin{aligned} g^{k+1}(V_a) &> \theta g^{k+1}(V^k) + (1-\theta)g^{k+1}(V^{k+1}) \\ &\overset{(a)}{>} \theta g^{k+1}(V^{k+1}) + (1-\theta)g^{k+1}(V^{k+1}) \\ &= g^{k+1}(V^{k+1}), \end{aligned} \tag{4.15}$$

where (a) follows from the proof of Lemma 4.3.1. Moreover,

$$\begin{aligned} g^{k+1}(V_a) &< g^{k+1}(V^k) + \text{tr}\left[\nabla g^{k+1}(V^k)^\top (V_a - V^k)\right] \\ &\overset{(a)}{=} g^{k+1}(V^k) + (1-\theta)\text{tr}(\nabla g^{k+1}(V^k)^\top (V^{k+1} - V^k)) \\ &\overset{(b)}{\leq} g^{k+1}(V^k), \end{aligned} \tag{4.16}$$

where (a) follows from linearity of the trace operator and (b) from optimality of V^{k+1} according to (4.12). Inequalities (4.15) and (4.16), imply that

$$g^{k+1}(V^k) > g^{k+1}(\theta V^k + (1-\theta)V^{k+1}) > g^{k+1}(V^{k+1}),$$

for all $\theta \in [0,1]$ – from which it follows that the sequence $\{V^k\}$ is strongly downward in the function g^{k+1} for each k. ■

We now prove Theorem 4.3.3 (a). The interested reader is refered to Theorem 1.4.3 in [62] for more details. Suppose, for the sake of contradiction, that $\lim_{k\to\infty} \|V^k - V^{k+1}\|_F \geq \varepsilon > 0$. Without loss of generality, consider two subsequences, such that $\{V^{k_n}\}_n \to \overline{X}$ and $\{V^{k_n+1}\}_n \to \widehat{V}$. By assumption, for every $\varepsilon > 0$,

$$\|V^{k_n+1} - V^{k_n}\|_F \geq \varepsilon > 0, \quad \text{for all } n \geq 1.$$

Because $\mathcal{C}$ is closed it contains all its limit points. Therefore,

$$\|\overline{V} - \widehat{V}\|_F \geq \varepsilon > 0.$$

The sequence $\{g^k(V^k)\}$ is monotonic non-increasing (Lemma 4.3.1) and since g^k is continuous on a compact set, g^k is bounded from below for all k. It follows from the monotone convergence theorem that the sequence $\{g^k(V^k)\}$ converges, i.e.

$$\lim_{k\to\infty} (g^{k+1}(V^{k+1}) - g^k(V^k)) = 0.$$

It follows that $g(\widehat{V}) = g(\overline{V})$. And by convexity of $\mathcal{C}$, we have that $\theta\overline{V} + (1-\theta)\widehat{V} \in \mathcal{C}$ for all $\theta \in [0,1]$. Moreover, the sequence $\{V^k\}$ is strongly downward in g^{k+1}. Therefore,

$$g^{k+1}(V^{k+1}) \leq g^{k+1}(\theta V^k + (1-\theta)V^{k+1}) \leq g^{k+1}(V^{k+1}) \overset{(a)}{<} g^k(V^k),$$

where (a) follows from Lemma 4.3.1. Taking limits gives,

$$g(\overline{V}) = g(\theta\overline{V} + (1-\theta)\widehat{V}) = g(\widehat{V}),$$

which contradicts the fact that g is hemivariate (Lemma 4.3.6). Therefore, $\lim_{k\to\infty}(V^k - V^{k+1}) = 0$. This completes the proof of (a).

We now prove part (b) of Theorem 4.3.3. Let $\overline{V} = \lim_{k\to\infty} V^k$. Since V^{k+1} was chosen to minimize 4.12, it must be true that

$$\operatorname{tr}\left(\nabla g^{k+1}(V^k)^*(V^{k+1} - V^k)\right) \leq \operatorname{tr}\left(\nabla g^{k+1}(V^k)^*(V - V^k)\right),$$

for all $V \in \mathcal{C}$. Taking limits and applying Lemma 4.3.1 yields

$$0 \leq \text{tr}\left(\nabla g(\overline{V})^*(V - \overline{V})\right).$$

Since the limit point was chosen arbitrarily, it follows from definition (4.3.2) that every limit point of $\{V^k\}$ is a stationary point. This complete the proof of part (b). ∎

While the proof of Theorem 4.3.3 follows largely from arguments in [62], it is included for completeness, as it ameliorates a minor gap in the proof of a similar result (Theorem II.2) appearing in [55].

4.3.2 Rank Minimization Heuristic

In a similar spirit with previous work [22, 55], we now consider a surrogate family for rank of the *log-det* type. More precisely, we define the sequence of surrogates $\{g^k\}$ as

$$g^k(V) = \log\det(f(V) + \delta^k I), \quad k = 1, 2, \dots \tag{4.17}$$

where the sequence of regularization parameters $\{\delta^k\}$ is assumed to be monotonic non-increasing with a finite limit $\delta > 0$. Moreover, we restrict $f : \mathbf{H}_+^n \to \mathbf{H}_+^n$ to a family of mappings that preserve *strict concavity* and *continuous differentiability* of g^k on a convex, compact subset $\mathcal{C}$ of the positive semidefinite cone for all k. It follows readily, by the monotonic convergence of $\{\delta^k\} \to \delta > 0$, that $\{g^k\}$ is a monotonic sequence of functions satisfying

$$\lim_{k\to\infty} g^k(V) = g(V) := \log\det(f(V) + \delta I)$$

for every in $V \in \mathcal{C}$. The gradient is easily computed as

$$\nabla g^k(V) = (f(V) + \delta^k I)^{-1} \nabla f(V).$$

An iterative linearization-minimization of the functions $\{g^k\}$ in (4.17) will converge, by Theorem 4.3.3, to a stationary point of g on some compact set. Section 4.3.4 discusses the selection of regularization coefficients $\{\delta^k\}$ to ensure certain rank monotonicty properties of the iterates $\{V^k\}$.

Remark 9. *For the identity mapping $f(V) = V$, we recover the classical log-det heuristic [22]. Other natural candidates for f include the* quadratic*, $f(V) = V^*V$, or* exponential mappings*, $f(V) = I - \exp(-\tau V)$, where $\tau > 0$ is a regularization constant controlling the concavity of f. We remark that there exists a broad literature quantifying, both analytically and empirically, the behavior of a much larger family of rank surrogates that go beyond the log-det family. However, such a discussion is beyond the scope of this paper and we refer the reader to [66, 56, 78] for a partial cross section of relevant literature.*

Recall from Section 3.2.2, our objective of efficiently extracting hidden rank one matrices belonging to the optimal face $\mathcal{F}^\star$ of the semidefinite relaxation of the AC-OPF problem. Leveraging on the preceding development, we now offer a simple iterative heuristic in Table 4.3 with the aim of doing precisely that. Given a high rank (>1) solution $V^\star \in \mathcal{F}^\star$ to the semidefinite relaxation, we initialize the iterative linearization-minimization algorithm with a feasible set restricted to the optimal face $\mathcal{F}^\star$, and initial condition $V^\star$. For notational brevity, we denote the iterative linearization-minimization algorithm in Table 4.3 as the mapping

$$\overline{V} \ = \ \Gamma(\mathcal{F}^\star, V^\star),$$

where $\overline{V} \in \mathcal{F}^\star$ denotes the converged value (within a prescribed tolerance) of the gradient descent method.

Algorithm 1: $\overline{V} = \Gamma(\mathcal{C}, V^0)$

Given a convex, compact set $\mathcal{C} \subset \mathcal{D}$, an initial condition V^0, a stopping tolerance $\varepsilon > 0$, and maximum number of iterations $\overline{k}$
Initialize $k = 0$
Repeat

1. *Compute.* $V^{k+1} \in \underset{V \in \mathcal{C}}{\text{argmin}} \ \text{tr}(\nabla g^{k+1}(V^k)^\top V)$

2. *Update.* $k = k + 1$

Until $\|V^k - V^{k-1}\|_F < \varepsilon$ or $k = \overline{k}$
Output $\overline{V} = V^k$

Table 4.3: Iterative Linearization-Minimization Algorithm

4.3.3 Alternating Bisection-Minimization Algorithm

In the event that the rank minimization heuristic fails to yield a rank one solution in $\mathcal{F}^\star$ (i.e., $\text{rank}(\overline{V}) > 1$), one of two motives could be at play. Firstly, there may not exist a rank one point belonging to the optimal face $\mathcal{F}^\star$ (cf. category C2). Secondly, while there may exist a rank one point in $\mathcal{F}^\star$, the heuristic may fail to recover it, as we have provided no guarantee on the algorithm's ability to recover a minimum rank solution. In either case, we offer in Table 4.4 a simple bisection method to iteratively relax the set of feasible points until a rank one feasible point is obtained. And naturally, there is no guarantee as to whether the resulting rank one point is globally optimal for the original AC-OPF problem (4.3), unless the global lower bound $J^\star$ is achieved.

The iterative relaxation of the feasible set obeys a simple *bisection rule* described as follows. First, let $\overline{J}$ denote a global upper bound on the optimal cost of the AC-OPF problem – a quantity that most commercial solvers can readily provide. If the rank minimization heuristic (cf. Table 4.3) fails to recover a rank one point on the optimal face $\mathcal{F}^\star$, i.e.

$$\operatorname{rank}(\overline{V}^0) > 1, \quad \text{where } \overline{V}^0 = \Gamma(\mathcal{F}^\star, V^\star),$$

we enlarge the feasible set to include points incurring a cost no greater than than the bisection point, $J^1 := J^\star + 0.5(\overline{J} - J^\star)$ in the interval $[J^\star, \overline{J}]$. The expanded feasible set is

$$\mathcal{F}^1 = \{V \in \mathcal{D} \;:\; J(V) \leq J^1\},$$

and apply the rank minimization heuristic over the new initial condition $\overline{V}^0$ and feasible set $\mathcal{F}^1$ to obtain an updated solution $\overline{V}^1 = \Gamma(\mathcal{F}^1, \overline{V}^0)$. The subsequent decision to bisect from above or below J^1, at the following time step, depends on the rank of the current solution $\overline{V}^1$. This alternation between bisection and optimization repeats ad nauseum until the bisection points converge to within a prescribed tolerance of one another. We refer the reader to Table 4.4 for a precise description of said method.

Remark 10. *We mention two caveats. First, for certain realizations of the AC-OPF problem, one may not be able to efficiently obtain a global upper bound, $\overline{J}$, through which to parameterize the bisection algorithm, as finding a point belonging to the non-convex feasible set of AC-OPF is, in general, NP-hard. Second, the bisection algorithm's ability to recover a rank one solution may be sensitive to the recursive choice of initial condition for the rank minimization algorithm at each bisection step. We have suggested one possible recursion, where the solution at the previous bisection step, initializes the rank minimization algorithm at the current step. One can imagine many variations in said scheme.*

4.3.4 Rank Monotonicity

Success of the iterative rank minimization algorithm (4.3) hinges on its convergence to a rank one point belonging to the optimal face $\mathcal{F}^\star$. As such, it's natural to ask as to

Algorithm 2: *Alternating Bisection-Minimization*

Given bounds (ℓ^0, u^0), an initial condition $\overline{V}^0$, and stopping tolerance $\varepsilon > 0$

1. *Bisect*. $J^1 = \ell^0 + \frac{1}{2}(u^0 - \ell^0)$
2. *Set*. $k = 1$

Repeat

7. *Update set.* $\mathcal{F}^k = \{V \in \mathcal{D} : J(V) \leq J^k\}$
8. *Call Algorithm 1.* $\overline{V}^k = \Gamma(\mathcal{F}^k, \overline{V}^{k-1})$
9. **if** $\mathrm{rank}(\overline{V}^k) > 1$
   (a) *Bisect from above.* $J^{k+1} = J^k + \frac{1}{2}(u^k - J^k)$
   (b) *Update bounds.* $\ell^{k+1} = J^k,\ u^{k+1} = u^k,$
10. **else if** $\mathrm{rank}(\overline{V}^k) = 1$
    (a) *Bisect from below.* $J^{k+1} = \ell^k + \frac{1}{2}(J^k - \ell^k)$
    (b) *Update bounds.* $\ell^{k+1} = \ell^k,\ u^{k+1} = J^k,$
11. *Update time.* $k = k + 1$

Until $|J^k - J^{k-1}| < \varepsilon$
Output $J^k,\ \overline{V}^{k-1}$

Table 4.4: Alternating Bisection-Minimization Algorithm

whether the iterates $\{V^k\}$ are monotonic in rank? Namely, can one guarantee that the $\mathrm{rank}(V^{k+1}) \leq \mathrm{rank}(V^k)$ for all k? This is a nuanced question, as the practical evaluation of rank requires approximation.

The rank of a matrix is equal to the number of non-zero singular values of the matrix. This fact is useful for theoretical analyses but it raises subtle issues when performing numerical computations with floating point numbers. In particular, the finite precision of floating point arithmetic implies that nonzero singular values cannot be distinguished from zero if their magnitude is sufficiently small. Conversely, numerical errors that arise in floating point computations can cause a matrix to have spurious non-zero singular values. As a consequence a threshold tolerance is typically used to determine the number of non-zero singular values and hence the rank of a matrix. To be precise, the numerical results generated in this paper calculate the rank of a matrix as the number of singular values that exceed a certain threshold ε. These numerical issues related to the matrix rank raise interesting questions that should be addressed by any practical semidefinite programming algorithm.

In light of the preceding discussion, we introduce a notion of *near low rank*, which is meant to capture matrices that are well approximated by low rank matrices. More precisely, we have the following definition.

Definition 4.3.7. *A matrix $X \in \mathbf{C}^{n\times n}$ is defined to be* ε-near rank-p *if X satisfies*

$$X \;=\; M + N, \qquad M, N \in \mathbf{C}^{n\times n}$$

where $\operatorname{rank}(M) = p$ *and* $\|N\|_F \leq \varepsilon$.

Equivalently, a matrix is said to be ε-near rank-p if it lives within a ε-radius ball centered around a rank-p matrix.

We now explore certain rank monotonicity properties of the matrix iterates $\{V^k\}$ generated by the the rank minimization heuristic (4.12) under the sequence of surrogates $g^k(V) = \log\det(V + \delta^k I)$.

Theorem 4.3.8 (Near rank monotonicity)**.** *Let* $\operatorname{rank}(V^k) = p \geq 1$. *Then V^{k+1} is ε-near rank-r (where $r \leq p$), if*

$$\delta^{k+1} \;\leq\; \frac{\varepsilon}{p}.$$

Power System (n)	Reference	$\text{rank}(V^0)$	$\text{rank}(\overline{V}^0)$	Iteration	$J^\star$	$\overline{J}$
9	[71]	3	1	23	1458.8	1458.8
30	[2]	2	1	11	316.49	316.49

Table 4.5: Power system examples with hidden rank one optimal solutions.

Remark 11 (Approximate constraint satisfaction)**.** *Using this notion of near low rank, one can pose interesting questions regarding* approximate constraint satisfaction*. For example, consider an ε-near rank one matrix $X = M + N$ (where $\text{rank}(M) = 1$ and $\|N\|_F \leq \varepsilon$) belonging to the optimal face $\mathcal{F}^\star$ of the AC-OPF semidefinite relaxation. While the naive rank one approximation $X \approx M$ may result in a violation of constraints (i.e. $M \notin \mathcal{F}^\star$), the violation will be mild. And for practical engineering problems such as OPF, minor constraint violations may be tolerable. It's therefore natural to ask as to when the optimal face $\mathcal{F}^\star$ possesses nearly rank one matrices that can be efficiently computed? Conversely, for semidefinite relaxations which do not possess rank one optimal solutions, can one systematically and efficiently construct a* mild relaxation *of the optimal face $\mathcal{F}' \supset \mathcal{F}^\star$ such that $\mathcal{F}'$ admits a rank one matrix?*

4.4 Numerical Studies

The primary objective of this section is to present a cross section of numerical results on the performance of the linearization-minimization and alternating bisection-minimization algorithms. A number of representative power system examples are presented for which the semidefinite relaxation fails. In Section 4.4.1, examples are provided for which the linearization-minimization algorithm succeeds in finding *hidden* rank one optimal solutions that are also globally optimal for the original AC-OPF problem. In the event that said algorithm fails to find a rank one matrix on the optimal face, the alternating bisection-minimization method can be applied. In Section 4.4.2, this alternating bisection-minimization method is used to find a rank one feasible solution that yields a cost no larger than that obtained via a conventional nonconvex solver. Throughout this section V^0, $\overline{V}^0$ and $\overline{V}^k$ denote, respectively, the optimal solution to the semidefinite relaxation, the rank minimization heuristic over $\mathcal{F}$, and the k^{th} step of the alternating bisection-minimization method.

4.4.1 Linearization-Minimization Iteration

Table 4.5 summarizes the power system networks used to test the linearization-minimization algorithm. For each example, the semidefinite relaxation fails to return a rank one solution (column 2). In each case, the linearization-minimization algorithm successfully converges to a rank one optimal point (column 3) typically in a small number of iterations (column 4). Thus the optimal cost for the semidefinite relaxation, $J^{\star}$, is in fact equal to the optimal cost of the AC-OPF problem. Moreover, the rank one solution returned from the linearization-minimization algorithm can be used to construct an optimal solution for the AC-OPF problem. These results verify that primal/dual solvers will fail to return rank one optimal solutions for the semidefinite relaxation even when such solutions exist (cf. Theorem 4.1.1). The values of $\overline{J}$ in the last column denote the upper bound on the optimal cost of the AC-OPF problem given by the nonconvex solver Matpower [96].

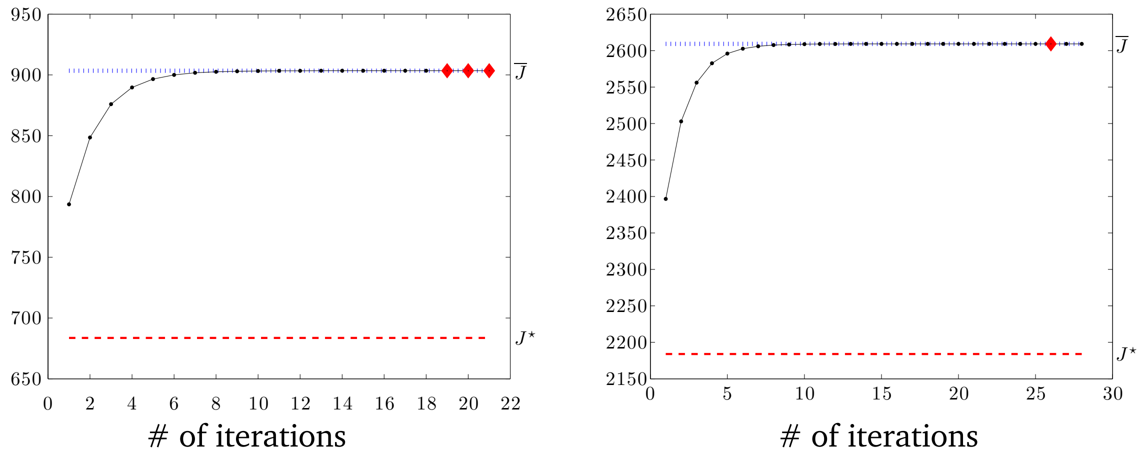
Power System (n)	Reference	$\mathrm{rank}(V^0)$	$\mathrm{rank}(\overline{V}^0)$	Iteration	$J^\star$	$\overline{J}$
3	[43]	2	2	1	683.62	944.34
5	[42]	3	2	6	2184.0	2609.3

Table 4.6: Power system cases for which the linearization-minimization algorithm fails to return a rank one matrix in the optimal solution set of the semidefinite relaxation.

4.4.2 Alternating-Bisection Algorithm

For certain problems, the linearization-minimization algorithm fails to return a rank one point in $\mathcal{F}^\star$ – i.e., $\mathrm{rank}(\overline{V}^0) > 1$. In such cases, one of two scenarios could be at play. Either the optimal face $\mathcal{F}^\star$ of the semidefinite relaxation does not possess a rank one matrix or the rank minimization heuristic may simply fail in returning a rank one point in $\mathcal{F}^\star$ when it does in fact exist. Table 4.6 provides three representative examples of such cases. For certain examples, the rank minimization heuristic is able to find a lower rank matrix (on $\mathcal{F}^\star$) than that achieved by the semidefinite relaxation. However, the iteration does not converge to a rank one solution. In each case, there is a nonzero gap between the cost achieved for the semidefinite relaxation, $J^\star$, and the Matpower upper bound obtained for the AC-OPF problem, $\overline{J}$.

The alternating bisection-minimization method is applied to the cases in Table 4.6. Figure 4.1 depicts the cost of a feasible point produced at every step of the bisection for the examples considered in Table 4.6. The red diamonds denote the iterates achieving rank one feasible points, while the black circles denote iterates corresponding to high rank feasible points. We observe in Figure 4.1, that in the case of the three and five bus examples, the minimum cost obtained by a rank one feasible point through bisection coincides with the cost produced by Matpower. This may lead one to believe that the optimal face $\mathcal{F}^\star$ of the semidefinite relaxation may not admit a rank one feasible point. In fact, for the three bus



(a) (b)

Figure 4.1: $J(\overline{V}^k)$ vs # of Iterations (Bisection Method) (a) 3-Bus Network (b) 5-Bus Network.

example, we can verify through Theorem 2.5.8 that the optimal solution of the semdefinite relaxation is unique and has rank greater than one.

To summarize, we observe that in many cases the iterative linearization-minimization algorithm successfully uncovers a hidden rank one point that is also globally optimal for the original AC-OPF problem. If the rank minimization algorithm fails to uncover a rank one optimal point, then the alternating bisection-minimization method can be applied. In this case, a rank one feasible solution is obtained that yields a cost that is no greater than that achieved by Matpower.

Part III

Robust AC Optimal Power Flow

CHAPTER 5

ROBUST AC OPTIMAL POWER FLOW

5.1 Introduction

The AC optimal power flow (AC-OPF) problem is a fundamental decision problem in power system operations [17]. In its most general form, AC-OPF amounts to a nonconvex optimization problem, where the objective is to minimize the total cost of generation subject to power balance constraints described by Kirchhoffs current and voltage laws, and operational constraints reflecting real and reactive limits on power generation, branch flows, and bus voltage magnitudes. The nonconvexity of AC-OPF derives in part from the need to enforce quadratic constraints, which are indefinite in the vector of bus voltages. The treatment of such nonconvexities in the AC-OPF problem has traditionally relied on the use of local methods for constrained optimization, or the use of approximate linear models of power flow to convexify the feasible set of the underlying optimization problem, e.g., DC-OPF [81]. More recently, considerable effort has been made to identify conditions under which an optimal solution to AC-OPF can be obtained from a solution to its semidefinite programming relaxation [41, 48, 49, 81].

Increased environmental concerns surrounding climate change have induced many U.S. states to adopt legislation mandating that a significant percentage of their electricity be generated by clean renewable resources. A basic challenge facing the large-scale integration of wind and solar resources derives from the need to compensate the attending intermittency and uncertainty in their supply of power. At the heart of this challenge is the need to develop robust optimization methods for AC-OPF to enable the reliable and cost-effective operation of power systems with a large fraction of their power supplied from uncertain renewable sources. In its most basic form, the *robust AC optimal power flow* (RAC-OPF) problem amounts to a two-stage robust optimization problem, in which the system operator must determine a day-ahead generation schedule, which minimizes the expected cost of dispatch, given a recourse opportunity to adjust its day-ahead schedule in real-time when the uncertain system variables have been realized, e.g., the available supply from renewable resources. The need to optimize over (infinite-dimensional) recourse policies, coupled with the intrinsic nonconvexity of the AC power flow constraints, results in RAC-OPF being an infinite-dimensional, nonconvex optimization problem in its most general form.

To address the nonconvexity in RAC-OPF, the vast majority of the literature on the topic employs a DC linear approximation of the power flow model [10, 35, 36, 44, 45, 59, 69, 68, 79, 87]. In addition to this approximation, the majority of the literature relies on affine or piecewise-affine approximations of the infinite-dimensional recourse policy space [3, 10, 35, 36, 44, 45, 59, 69, 68, 79, 85, 87]. The primary approach to the treatment of uncertainty has focused on either *robust* [3, 35, 36, 59, 63, 87] or *chance-constrained* [10, 79, 69, 68, 85] formulations.

Contribution: In this chapter, we formulate RAC-OPF as a two-stage robust optimization problem with recourse. Our primary point of departure from the existing literature is our treatment of the full AC power flow model. To the best of our knowledge, the only papers that treat the AC power flow model are [3, 85, 63]. However, a critical assumption made in these papers is the assumption of exactness of the convex (semidefinite or second-order cone) programming relaxations on which they rely. Exactness of such convex relaxations for RAC-OPF is not guaranteed, and, in particular, the solutions generated by these relaxations are not guaranteed to be feasible for RAC-OPF. In this chapter, we restrict the space of recourse policies to those which are *affine* in the uncertain problem variables, and provide a method to approximate the RAC-OPF problem from within by a finite-dimensional semidefinite program. The resulting semidefinite program – a convex inner approximation to RAC-OPF – produces affine recourse policies that are *guaranteed to be feasible* for RAC-OPF.

In addition to the inner approximation, we we develop a method for constructing an outer approximation (relaxation) to the RAC-OPF problem as a finite-dimensional second-order cone program. Our approach centers on the reformulation of the RAC-OPF problem as a robust rank one constrained semidefinite program and on its relaxation to a robust linear program. This relaxation is obtained by eliminating the rank constraint and by approximating the cone of positive semidefinite matrices from without by a polyhedral cone. As the proposed relaxation depends on the choice of the polyhedral cone, we propose a a recursive method, which refines said cone in a manner guided by the objective function to improve the performance of the relaxation. The practical value of our approximation techniques proposed in this chapter derive from the fact that one can obtain a feasible solution to the RAC-OPF problem by solving a finite-dimensional semidefinite program; and can bound the suboptimality incurred by this feasible solution by solving another finite-dimensional conic linear program. And if the gap between the optimal values of the outer and inner approximations is small, we have a certificate of near optimality of the feasible solution.

The chapter is organized as follows. In Sections 5.2 and 5.3, we develop the power system model and provide a detailed formulation of RAC-OPF, respectively. In Section 5.4, we offer a detailed derivation of the semidefinite programming inner approximation of

RAC-OPF, and provide a sufficient condition under which the resulting approximation is guaranteed to have a nonempty feasible region. In Section 5.4.4, we describe an iterative optimization method that generates a sequence of feasible affine recourse policies with nonincreasing costs. In Section 5.5, we derive the second-order cone outer approximation (relaxation) of RAC-OPF. We offer an recursive method for improving the effectiveness of the relaxation in Section 5.5.2. Finally, we illustrate the effectiveness of the proposed approximation methods on a nine-bus power system with different levels of renewable resource penetration and uncertainty in Section 5.6.

Notation: Let $\mathbf{N},\ \mathbf{C},$ and $\mathbf{R}$ be the sets of natural, complex, and real numbers, respectively. For any $m \in \mathbf{N}$, let $[m] := \{1, 2, \ldots, m\}$. Denote by e_i the real i^{th} standard basis vector, of dimension appropriate to the context in which it is used. For any $z_1, z_2 \in \mathbf{C}$, we define a partial ordering on $\mathbf{C}$ by $z_1 \leq z_2$ if and only if $\mathrm{Re}\{z_1\} \leq \mathrm{Re}\{z_2\}$ and $\mathrm{Im}\{z_1\} \leq \mathrm{Im}\{z_2\}$. For any X, let $[X]_{ij}$ denote its (i, j) entry, and X^* its conjugate transpose. Let $\mathbf{H}^n$ be the space of n-by-n Hermitian matrices. For a matrix $X \in \mathbf{H}^n$, the notation $X \succeq 0$ means that X is positive semidefinite. For a matrix $X \in \mathbf{R}^{m \times n}$, $X \geq 0$ means that X is entrywise nonnegative. Let $\mathrm{tr}(X)$ denote the trace of a matrix X. Denote by I_n the n-by-n identity matrix. Finally, for any $k, n \in \mathbf{N}$, let $\mathbf{L}^2_{k,n}$ $(\mathbf{L}^2_{k,n\times n})$ be the space of all Borel measurable, square-integrable functions from $\mathbf{R}^k$ to $\mathbf{C}^n$ $(\mathbf{H}^n)$.[1]

[1]A complex-valued function f on $\mathbf{R}^k$ is said to be Borel measurable if both $\mathrm{Re}\{f\}$ and $\mathrm{Im}\{f\}$ are real-valued Borel measurable.

5.2 Power System Model

We begin with a development of a general model for AC optimal power flow under uncertainty. The power system we consider consists of a heterogeneous mix of generation and load resources, which differ in terms of their inherent controllability and predictability. The perspective we adopt is that of the inpdendent system operator (ISO), whose objective is to determine the dispatch of available generation resources in order to minimize the expected cost of meeting demand, while ensuring that all operational limits of generation and transmission facilities are met. This problem is commonly referred to as the *security constrained economic dispatch* (SCED) problem. The optimization model we consider consists of two stages: day-ahead (DA) and real-time (RT). In day-ahead, the ISO must schedule an initial dispatch of its resources subject to uncertainty in the eventual realization of certain system variables, e.g., demand and available supply of renewable resources. Such DA scheduling decisions are essential, as certain generation resources (e.g., coal and nuclear) have limited ramping capability, and must therefore be scheduled to produce well in advance of the delivery time. In real-time, all uncertain variables are realized, and the ISO is provided a recourse opportunity to adjust its DA dispatch schedule in order to balance the system at minimum cost. Ultimately, the determination of a DA schedule, which minimizes the expected cost of dispatch given optimal recourse in real-time requires the solution of a two-stage robust optimization problem with recourse. We formally define this problem in (5.8).

5.2.1 Power Flow Model

We consider a power transmission network whose topology is described by an undirected graph $\mathcal{G} := (\mathcal{V}, \mathcal{E})$. Here, the vertex set $\mathcal{V} := [n]$ represents the collection of transmission buses, and the edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ represents the collection of transmission lines connecting buses. We require that $(i, j) \in \mathcal{E}$, if and only if $(j, i) \in \mathcal{E}$.

We describe the *AC power flow* equations according to Kirchhoff's current and voltage laws, which govern the relationship between complex bus voltages and power injections [9]. Let $Y \in \mathbf{C}^{n \times n}$ denote the network admittance matrix, $v \in \mathbf{C}^n$ the vector of complex bus voltages, and $s \in \mathbf{C}^n$ the vector of complex (net) bus power injections (generation minus demand). The AC power balance equations can be expressed as

$$s_i = v^* S_i v, \tag{5.1}$$

where $S_i := Y^* e_i e_i^*$ for all $i \in \mathcal{V}$. For each line $(i, j) \in \mathcal{E}$, we denote the complex power flow from bus i to bus j by $s_{ij} \in \mathbf{C}$. It is given by

$$s_{ij} = v^* S_{ij} v, \tag{5.2}$$

where $S_{ij} := e_i e_i^* (\widehat{y}_{ij}/2 - [Y]_{ij})^* + e_j e_i^* [Y]_{ij}^*$ for all $(i, j) \in \mathcal{E}$. Here, $\widehat{y}_{ij} \in \mathbf{C}$ denotes the total shunt admittance of line (i, j).

We require that the following constraints be enforced. The first class of constraints requires that bus voltage magnitudes satisfy

$$v_i^{\min} \leq |v_i| \leq \ v_i^{\max},$$

for all buses $i \in \mathcal{V}$. Here, $v_i^{\min} \in \mathbf{R}$ and $v_i^{\max} \in \mathbf{R}$ denote upper and lower bounds, respectively, on the voltage magnitude at bus i. The second class of constraints we consider enforce power flow capacities on the transmission lines. Namely, for each transmission line $(i,j) \in \mathcal{E}$, the real power flow from bus i to bus j must satisfy

$$-\ell_{ij}^{\max} \leq v^* P_{ij} v \leq \ell_{ij}^{\max}.$$

Here, $P_{ij} := (S_{ij} + S_{ij}^*)/2$, and $\ell_{ij}^{\max} \in \mathbf{R}$ denotes the real power flow capacity of line (i,j).

We refer the reader to Appendix A and to the references therein for a detailed derivation of the power flow and power balance equations.

5.2.2 Uncertainty Model

All of the 'uncertain' quantities appearing in this chapter are described according to the random vector $\boldsymbol{\xi}$, which is defined according to the probability space $(\mathbf{R}^k, \mathcal{B}(\mathbf{R}^k), \Pr)$. Here, the Borel σ-algebra $\mathcal{B}(\mathbf{R}^k)$ is the set of all events that are assigned probabilities by the measure $\Pr$. We denote the first and second-order moments of $\boldsymbol{\xi}$ by

$$\mu := \mathbb{E}[\boldsymbol{\xi}] \text{ and } M := \mathbb{E}[\boldsymbol{\xi}\boldsymbol{\xi}^*],$$

where $\mathbb{E}[\cdot]$ denotes the expectation operator with respect to $\Pr$. Adopting a standard notational convention, we will use ξ (normal face) to denote realizations taken by the random vector $\boldsymbol{\xi}$ (bold face). We assume throughout the paper that the support of the random vector $\boldsymbol{\xi}$ is nonempty, compact, and representable as

$$\Xi := \{\xi \in \mathbf{R}^k \mid \xi_1 = 1,\ \xi^* W_j \xi \geq 0,\ j = 1, \ldots, \ell\}. \tag{5.3}$$

Here, each matrix $W_j \in \mathbf{R}^{k \times k}$ is defined according to

$$W_j := \begin{bmatrix} \omega_j & w_j^* \\ w_j & -\Omega_j^* \Omega_j \end{bmatrix}, \tag{5.4}$$

where $\omega_j \in \mathbf{R}$, $w_j \in \mathbf{R}^{k-1}$, and $\Omega_j \in \mathbf{R}^{n_j \times (k-1)}$ for some $n_j \in \mathbf{N}$. It is important to note that the representation of the support in (5.3) is general enough to describe any subset of the hyperplane $\{\xi \in \mathbf{R}^k \mid \xi_1 = 1\}$ that is defined according to a finite intersection of arbitrary half spaces and ellipsoids. We will occasionally refer to the support set Ξ as the 'uncertainty set' associated with the random vector $\boldsymbol{\xi}$.

Some remarks regarding our uncertainty model are in order. First, the requirement that $\xi_1 = 1$ for all $\xi \in \Xi$ is for notational convenience, as it allows one to represent affine functions of $(\xi_2, \ldots, \xi_k)$ as linear functions of ξ. Second, it is important to emphasize that all of the results contained in this chapter depend on the probability distribution of the random vector $\boldsymbol{\xi}$ only through its support, mean, and second-order moment. No additional information about the distribution is required. We make the following mild technical assumption that is assumed to hold throughout the paper.

Assumption 5.2.1. *There exists $\xi \in \Xi$ such that $\xi^* W_j \xi > 0$ for all $j \in [\ell]$.*

The assumption that the support set Ξ admits a strictly feasible point will prove useful to the derivation of our subsequent theoretical results, as it ensures that Ξ spans all of $\mathbf{R}^k$. This, in turn, guarantees that the second-order moment matrix M is positive definite and invertible. We refer the reader to [40, Prop. 2] for a proof of this claim.

5.2.3 Generator and Load Models

Load Model: We consider a load model in which the real-time demand for power at each bus $i \in \mathcal{V}$ is fixed, known, and denoted by $d_i \in \mathbf{C}$.

Generator Model: To maintain clarity of exposition throughout the paper, we assume that there is at most a single generator at each bus $i \in \mathcal{V}$. We consider a generator model in which the real-time (RT) supply of power, as determined by the ISO, is allowed to depend on the realization of the random vector $\boldsymbol{\xi}$. Accordingly, we let $g_i(\boldsymbol{\xi})$ denote the power produced at bus i in real-time, where $g_i \in \mathbf{L}^2_{k,1}$ is a *recourse function* determined by the ISO for each bus $i \in \mathcal{V}$. Each generator i incurs a cost for producing $g_i(\boldsymbol{\xi})$, which we assume to be linear in the real power produced. We explicitly define its production cost as

$$\alpha_i \mathrm{Re}\{g_i(\boldsymbol{\xi})\}, \qquad i \in \mathcal{V}.$$

Here, $\alpha_i \geq 0$ denotes the marginal cost of real power generation at bus $i \in \mathcal{V}$.

In order to capture the potential for uncertainty in the generating capacity available to each generator in real-time, we require that the power produced by each generator respects the generation capacity constraints

$$\underline{g}_i(\boldsymbol{\xi}) \leq g_i(\boldsymbol{\xi}) \leq \overline{g}_i(\boldsymbol{\xi}), \qquad i \in \mathcal{V}. \tag{5.5}$$

Here, $\underline{g}_i(\boldsymbol{\xi}) \in \mathbf{L}^2_{k,1}$ and $\overline{g}_i(\boldsymbol{\xi}) \in \mathbf{L}^2_{k,1}$ denote the minimum and maximum power levels, respectively, that generator i can sustain in real-time. This uncertainty in a generator's available capacity can be used to model unscheduled generator outages, and intermittency in renewable power supply. The random generation capacities are assumed to satisfy

$$g_i^{\min} \leq \underline{g}_i(\boldsymbol{\xi}) \leq \overline{g}_i(\boldsymbol{\xi}) \leq g_i^{\max}, \qquad i \in \mathcal{V}.$$

Here, $g_i^{\min} \in \mathbf{C}$ and $g_i^{\max} \in \mathbf{C}$ are the *nameplate* minimum and maximum capacities of generator i, respectively. We denote the corresponding vectors by $\underline{g}(\boldsymbol{\xi})$, $\overline{g}(\boldsymbol{\xi})$, $g^{\min}$, $g^{\max}$.

In practice, a generator cannot adjust its production level instantaneously, but rather is limited by a prespecified rate (usually measured in MVA/min) that depends on the type of generator. We specify generator i's limited ramping capability in real-time according to the following pair of constraints

$$r_i^{\min} \leq g_i(\boldsymbol{\xi}) - g_i^0 \leq r_i^{\max}, \qquad i \in \mathcal{V}, \tag{5.6}$$

where $r_i^{\min} \in \mathbf{C}$ and $r_i^{\max} \in \mathbf{C}$ denote generator i's ramp-down and ramp-up limits, respectively. Here, $g_i^0 \in \mathbf{C}$ denotes generator i's day-ahead (DA) dispatch, also determined by the ISO. The DA dispatch of each generator is required to satisfy its nameplate generation capacity constraints given by

$$g_i^{\min} \leq g_i^0 \leq g_i^{\max}, \qquad i \in \mathcal{V}. \tag{5.7}$$

Example 5.2.2 (Generator types)**.** In line with [65], our generator model is general enough to capture a wide range of generator types. We specify several important examples in the following discussion. Given a DA dispatch level g_i^0 that satisfies (5.7), generator i is said to be:

- *Completely inflexible* (e.g., nuclear) if its allowable range of real-time outputs is given by

$$g_i(\boldsymbol{\xi}) = g_i^0.$$

- *Completely flexible* (e.g., oil, gas) if its allowable range of real-time outputs is given by

$$g_i^{\min} \leq g_i(\boldsymbol{\xi}) \leq g_i^{\max}.$$

- *Intermittent* (e.g., wind, solar) if its allowable range of real-time outputs is given by

$$\underline{g}_i(\boldsymbol{\xi}) \leq g_i(\boldsymbol{\xi}) \leq \overline{g}_i(\boldsymbol{\xi}).$$

We make the following technical assumption, which requires that the RT generation capacities exhibit a linear dependence on the random vector $\boldsymbol{\xi}$. Assumption 5.2.3 is assumed to hold throughout the paper.

Assumption 5.2.3. *There exist matrices $\underline{G} \in \mathbf{C}^{n\times k}$ and $\overline{G} \in \mathbf{C}^{n\times k}$ such that $\underline{g}(\boldsymbol{\xi}) = \underline{G}\boldsymbol{\xi}$ and $\overline{g}(\boldsymbol{\xi}) = \overline{G}\boldsymbol{\xi}$.*

5.3 Formulation of Robust AC-OPF

Building on the previously defined models, we formulate the *robust AC optimal power flow* (RAC-OPF) problem as follows.

$$\begin{aligned}
\text{minimize} \quad & \mathbb{E}\left[\sum_{i=1}^{n} \alpha_i \mathrm{Re}\{g_i(\boldsymbol{\xi})\}\right] \qquad (5.8)\\
\text{subject to} \quad & g^0 \in \mathbf{C}^n, \quad g \in \mathbf{L}^2_{k,n}, \quad v \in \mathbf{L}^2_{k,n} \\
& g_i^{\min} \leq g_i^0 \leq g_i^{\max}, \qquad i \in \mathcal{V} \\
& \left.\begin{aligned}
& \underline{g}_i(\boldsymbol{\xi}) \leq g_i(\boldsymbol{\xi}) \leq \overline{g}_i(\boldsymbol{\xi}), && i \in \mathcal{V} \\
& r_i^{\min} \leq g_i(\boldsymbol{\xi}) - g_i^0 \leq r_i^{\max}, && i \in \mathcal{V} \\
& g_i(\boldsymbol{\xi}) - v(\boldsymbol{\xi})^* S_i v(\boldsymbol{\xi}) = d_i, && i \in \mathcal{V} \\
& v_i^{\min} \leq |v_i(\boldsymbol{\xi})| \leq v_i^{\max}, && i \in \mathcal{V} \\
& |v(\boldsymbol{\xi})^* P_{ij} v(\boldsymbol{\xi})| \leq \ell_{ij}^{\max}, && (i,j) \in \mathcal{E}
\end{aligned}\right\} \ \mathbb{P}\text{-a.s.}
\end{aligned}$$

As previously described, the RAC-OPF problem amounts to a two-stage robust optimization problem with recourse. The single-period formulation of RAC-OPF that we consider is similar in structure to the single-period formulations studied in [3, 35, 85]. We briefly summarize the timing and structure of the decision variables and constraints of the RAC-OPF problem.

- The *first-stage* (day-ahead) decisions entail the determination of a DA generator dispatch $g_0 \in \mathbf{C}^n$ subject to optimal recourse in the second stage, which will adjust the DA dispatch given a realization of the random vector $\boldsymbol{\xi}$.
- In the *second stage* (real-time), the random vector $\boldsymbol{\xi}$ is realized, and the ISO is provided a recourse opportunity to adjust its DA generator dispatch in order to balance the system at minimum cost. The second-stage decision entails the determination of the RT generator dispatch $g(\boldsymbol{\xi}) \in \mathbf{L}^2_{k,n}$ and the RT bus voltages $v(\boldsymbol{\xi}) \in \mathbf{L}^2_{k,n}$.
- All decisions must be jointly determined in such a manner as to (i) minimize the expected cost of generation, and (ii) guarantee that all system constraints are satisfied given *any* realization $\xi \in \Xi$ of the random vector $\boldsymbol{\xi}$ in real-time, i.e., robust constraint satisfaction.

5.3.1 Concise Formulation

It will be convenient to our analysis in the sequel to work with a more concise representation of the RAC-OPF problem. We do so by first eliminating the RT generator dispatch variables $g \in \mathbf{L}^2_{k,n}$ through their direct substitution according to the nodal power balance equations. Second, by redefining the DA generator dispatch $g_0 \in \mathbf{C}^n$ as a real vector

$x := [\mathrm{Re}\{g_0\}^*, \mathrm{Im}\{g_0\}^*]^*$, one can rewrite problem (5.8) more compactly in the following form:

$$\begin{aligned}
\text{minimize} \quad & \mathbb{E}\left[v(\boldsymbol{\xi})^* A_0 v(\boldsymbol{\xi})\right] \\
\text{subject to} \quad & x \in \mathbf{R}^{2n},\ v \in \mathbf{L}^2_{k,n} \\
& v(\boldsymbol{\xi})^* A_i v(\boldsymbol{\xi}) + b_i^* x \le c_i^* \boldsymbol{\xi}, \quad i = 1, \ldots, m, \quad \mathbb{P}\text{-a.s.}, \\
& Ex \le f,
\end{aligned} \tag{$\mathcal{P}$}$$

where $m := 10n + 2|\mathcal{E}|$. It is straightforward to construct the matrices $E \in \mathbf{R}^{4n \times 2n}$, $f \in \mathbf{R}^{4n}$, $A_i \in \mathbf{H}^n$ $(i = 0, \ldots, m)$, $b_i \in \mathbf{R}^{2n}$ $(i = 1, \ldots, m)$, and $c_i \in \mathbf{R}^k$ $(i = 1, \ldots, m)$ given the underlying problem data specified in the RAC-OPF problem (5.8). We refer the reader to Appendix A.3 for their specification.

5.4 Convex Inner Approximation

Problem $\mathcal{P}$ is computationally intractable, in general, as it is both *infinite-dimensional* and *nonconvex*. The nonconvexity is due, in part, to the feasible set, which is defined by a number of indefinite quadratic constraints in the vector of complex bus voltages. The infinite-dimensionality of the optimization problem $\mathcal{P}$ derives from both the infinite-dimensionality of the recourse decision variables. In what follows, we develop a systematic approach to approximate problem $\mathcal{P}$ from within by a finite-dimensional semidefinite program, and provide a sufficient condition under which the resulting inner approximation is guaranteed to have a nonempty feasible region. The proposed method for approximation centers on the restriction of the infinite-dimensional space of recourse policies to those which are *linear* in the random vector $\boldsymbol{\xi}$.

5.4.1 Affine Recourse Policies

As the initial step in the derivation of a tractable inner approximation to problem $\mathcal{P}$, we first restrict the functional form of the recourse decision variables (i.e., the complex bus voltages) to be linear in the random vector $\boldsymbol{\xi}$. That is to say, we require that

$$v(\boldsymbol{\xi}) = V\boldsymbol{\xi}, \tag{5.9}$$

where $V \in \mathbf{C}^{n\times k}$. This restriction to affine recourse policies gives rise to the following optimization problem $\mathcal{P}_{\mathrm{I}}$, which stands as an inner approximation to the original problem $\mathcal{P}$.

$$\begin{aligned} \text{minimize} \quad & \mathrm{tr}(MV^*A_0V) \\ \text{subject to} \quad & x \in \mathbf{R}^{2n},\ V \in \mathbf{C}^{n\times k} \\ & \xi^*V^*A_iV\xi + b_i^*x \leq c_i^*\xi, \quad i = 1,\dots,m, \quad \forall\, \xi \in \Xi, \\ & Ex \leq f, \end{aligned} \tag{$\mathcal{P}_{\mathrm{I}}$}$$

We have used linearity of expectation and trace operators, and the invariance of trace under cyclic permutations to massage the original objective function to obtain $\mathbb{E}[\boldsymbol{\xi}^*V^*A_0V\boldsymbol{\xi}] = \mathbb{E}[\mathrm{tr}(\boldsymbol{\xi}\boldsymbol{\xi}^*V^*A_0V)] = \mathrm{tr}(\mathbb{E}[\boldsymbol{\xi}\boldsymbol{\xi}^*]V^*A_0V)$. We have also replaced the m almost sure constraints with robust constraints due to the continuity of the corresponding functions in ξ.

The resulting problem $\mathcal{P}_{\mathrm{I}}$ amounts to a semi-infinite nonconvex quadratically constrained quadratic program.[2] More specifically, the restriction to affine recourse policies results in an optimization problem $\mathcal{P}_{\mathrm{I}}$ whose decision variables range over finite-dimensional spaces.

[2]A semi-infinite program is an optimization problem involving finitely many decision variables, and an infinite number of constraints.

However, due to the continuous structure of the uncertainty set Ξ, problem $\mathcal{P}_{\mathrm{I}}$ has infinitely many constraints and is, in general, intractable. To account for this, we employ weak duality to obtain a *sufficient* set of finitely many constraints. We remark that such an approximation of the infinite constraint set can also be derived through a direct application of the so-called S-procedure [15]. We have the following result, which follows from Proposition 6 in [40].

Lemma 5.4.1. *Let $P \in \mathbf{H}^k$, $q \in \mathbf{R}^k$, $r \in \mathbf{R}$, and $Q := (e_1 q^* + q e_1^*)/2$. Consider the following two statements:*

(i) $\xi^* P \xi + q^* \xi + r \leq 0, \ \forall\, \xi \in \Xi$,

(ii) $\exists\, \lambda \in \mathbf{R}^\ell$ *with* $\lambda \leq 0$ *and* $P + Q + r e_1 e_1^* - \sum_{j=1}^{\ell} \lambda_j W_j \preceq 0$,

where W_j is as defined in (5.4). For any $\ell \in \mathbf{N}$, it holds that ((ii)) implies ((i)). If $\ell = 1$, then ((i)) and ((ii)) are equivalent.

Using Lemma 5.4.1, one can approximate the infinite constraint set of problem $\mathcal{P}_{\mathrm{I}}$ from within by finitely many matrix inequality constraints. More precisely, a direct application of Lemma 5.4.1 to each of the quadratic constraints in problem $\mathcal{P}_{\mathrm{I}}$ gives rise to the following finite-dimensional optimization problem:

$$
\begin{aligned}
\text{minimize} \quad & \mathrm{tr}(M V^* A_0 V) && (\mathcal{P}_{\mathrm{II}})\\
\text{subject to} \quad & x \in \mathbf{R}^{2n},\ V \in \mathbf{C}^{n\times k},\ \Lambda \in \mathbf{R}^{m\times \ell} \\
& V^* A_i V - C_i + (b_i^* x) e_1 e_1^* - \sum_{j=1}^{\ell} [\Lambda]_{ij} W_j \preceq 0, \\
& \forall\, i \in [m], \\
& Ex \leq f, \\
& \Lambda \leq 0,
\end{aligned}
$$

where we define $C_i := (e_1 c_i^* + c_i e_1^*)/2$ for each $i \in [m]$.

Remark 12. It follows directly from Lemma 5.4.1 that problem $\mathcal{P}_{\text{II}}$ is an inner approximation to problem $\mathcal{P}_{\text{I}}$ in general; and is equivalent to problem $\mathcal{P}_{\text{I}}$ when $\ell = 1$.

5.4.2 Convexifying the Inner Approximation

Problem $\mathcal{P}_{\text{II}}$ is a finite-dimensional inner approximation to the original problem $\mathcal{P}$. It, however, remains to be nonconvex, because of the indefinite quadratic functions appearing in both the objective and constraints. In what follows, we develop a method to convexify problem $\mathcal{P}_{\text{II}}$ from within by replacing each indefinite quadratic function with a *majorizing* convex quadratic function. We state the resulting convex program, which approximates $\mathcal{P}_{\text{II}}$ from within, in Proposition 5.4.3.

The proposed method is based on the simple observation that each indefinite quadratic function can be decomposed as a sum of a convex quadratic function and a concave quadratic function. We then approximate the concave function from above with its linearization at a point. The sum of this linearization with the convex component of the original function yields a convex global overestimator of the original indefinite quadratic function. More precisely, for each matrix A_i, define the decomposition

$$A_i = A_i^+ + A_i^-,$$

where $A_i^+ \succeq 0$ and $A_i^- \preceq 0$ denote the positive semidefinite and negative semidefinite parts of A_i, respectively. Using this matrix decomposition, define the function $H_i : \mathbf{C}^{n\times k} \times \mathbf{C}^{n\times k} \to \mathbf{H}^k$ according to

$$H_i(V, Z) := V^* A_i^+ V + Z^* A_i^- V + V^* A_i^- Z - Z^* A_i^- Z.$$

for each $i \in [m]$. The first term of H_i is the convex component of the original quadratic function $V^* A_i V$. The remaining terms represent the linearization of the concave component at a point Z. Consequently, for any matrix Z, the function $H_i(V, Z)$ is matrix convex in V.[3] The following result highlights two important properties of H_i. Its proof can be found in Appendix **??**.

Lemma 5.4.2. *Let $Z \in \mathbf{C}^{n\times k}$. For each $i \in [m]$, it holds that*

(i) $V^* A_i V \preceq H_i(V, Z), \quad \forall\, V \in \mathbf{C}^{n\times k}$*,*

(ii) $\mathrm{tr}(M V^* A_i V) \leq \mathrm{tr}(M H_i(V, Z)), \quad \forall\, V \in \mathbf{C}^{n\times k}$.

Proof. ((i)) For any $V, Z \in \mathbf{C}^{n\times k}$, consider the matrix inequality

$$0 \succeq (Z - V)^* A_i^- (Z - V) = V^* A_i V - H_i(V, Z), \tag{5.10}$$

where the first relation follows from the negative semidefiniteness of A_i^-. Rearranging, we obtain the desired matrix inequality.

((ii)) Since the matrix on the right-hand side of (5.10) is Hermitian negative semidefinite, we must have

$$N^*(V^* A_i V - H_i(V, Z))N \preceq 0,$$

where $N \in \mathbf{R}^{k\times k}$ is a Cholesky factor of M (i.e., $M = NN^*$). Such a matrix N is guaranteed to exist since M is positive definite. And since the trace of any Hermitian negative semidefinite matrix is nonpositive, we obtain

$$\mathrm{tr}(N^*(V^* A_i V - H_i(V, Z))N) \leq 0.$$

[3] A function $f : \mathbf{C}^{n\times k} \to \mathbf{H}^k$ is said to be matrix convex if for all matrices X, Y and $0 \leq \theta \leq 1$, we have $f(\theta X + (1-\theta)Y) \preceq \theta f(X) + (1-\theta) f(Y)$.

The linearity of the trace operator and its invariance under cyclic permutations give the desired trace inequality. ∎

Property ((i)) provides a way of approximating the nonconvex feasible set of problem $\mathcal{P}_{\mathrm{II}}$ from withing by a convex set. Property ((ii)), on the other hand, provides way of majorizing the nonconvex objective of problem $\mathcal{P}_{\mathrm{II}}$ with a convex function. In Proposition 5.4.3, we employ these approximations to specify a convex program whose optimal solution is guaranteed to be a feasible solution for the original problem $\mathcal{P}$. Its proof follows directly from Lemma 5.4.2. We, therefore, omit it for the sake of brevity.

Proposition 5.4.3. *Let $V_0 \in \mathbf{C}^{n\times k}$, and suppose that $(\overline{x}, \overline{V}, \overline{\Lambda})$ is an optimal solution for the following convex program:*

$$
\begin{aligned}
&\text{minimize} && \mathrm{tr}(MH_0(V, V_0)) && (\mathcal{P}_{\mathrm{III}}(V_0))\\
&\text{subject to} && x \in \mathbf{R}^p,\ V \in \mathbf{C}^{n\times k},\ \Lambda \in \mathbf{R}^{m\times \ell}\\
& && H_i(V, V_0) - C_i + (b_i^* x)e_1 e_1^* - \sum_{j=1}^{\ell} [\Lambda]_{ij} W_j \preceq 0,\\
& && \forall\, i \in [m],\\
& && Ex \le f,\\
& && \Lambda \le 0.
\end{aligned}
$$

Define the function $\overline{v} \in \mathcal{L}^2_{k,n}$ according to $\overline{v}(\boldsymbol{\xi}) = \overline{V}\boldsymbol{\xi}$. Then $(\overline{x}, \overline{v})$ is a feasible solution for the original problem $\mathcal{P}$.

We note that problem $\mathcal{P}_{\mathrm{III}}(V_0)$ can be equivalently reformulated as a semidefinite program using the Schur complement condition for positive semidefiniteness. We refer the reader to Appendix A.4 for the details of this reformulation.

5.4.3 Guaranteeing Nonemptiness of the Inner Approximation

In order to convexify the RAC-OPF problem according the method developed in Section 5.4.2, one has to select a matrix $V_0 \in \mathbf{C}^{n\times k}$ that results in a *nonempty feasible set* for the inner approximation $\mathcal{P}_{\text{III}}(V_0)$. In this section, we provide a method for computing one such matrix. The method we propose entails the calculation of a day-ahead dispatch $g^0 \in \mathbf{C}^n$, which is guaranteed to be feasible for the RAC-OPF problem without requiring adjustment (recourse) in real-time. In order to do so, one needs to first characterize the guaranteed range of available power supply at each bus in the network. For each bus $i \in \mathcal{V}$, this amounts to the specification of upper and lower limits $\gamma_i^{\max} \in \mathbf{C}$ and $\gamma_i^{\min} \in \mathbf{C}$, such that

$$\underline{g}_i(\xi) \le \gamma_i^{\min} \le \gamma_i^{\max} \le \overline{g}_i(\xi), \quad \forall\, \xi \in \Xi.$$

We calculate these limits according to

$$\begin{aligned} \gamma_i^{\min} &= \max_{\xi\in\Xi}\left(\text{Re}\{\underline{g}_i(\xi)\}\right) + \mathbf{j}\max_{\xi\in\Xi}\left(\text{Im}\{\underline{g}_i(\xi)\}\right), \\ \gamma_i^{\max} &= \min_{\xi\in\Xi}\left(\text{Re}\{\overline{g}_i(\xi)\}\right) + \mathbf{j}\min_{\xi\in\Xi}\left(\text{Im}\{\overline{g}_i(\xi)\}\right). \end{aligned} \tag{5.11}$$

It follows from Assumption 5.2.3 and the assumed structure of the uncertainty set Ξ that each of these quantities can be exactly calculated by solving a second-order cone program. A day-ahead dispatch, which is guaranteed to be feasible for the RAC-OPF problem, can therefore be calculated by solving the following deterministic AC-OPF problem.

$$\begin{aligned}
&\text{minimize} \quad \sum_{i=1}^{n} \alpha_i v^* \left(\frac{S_i + S_i^*}{2}\right) v && (5.12)\\
&\text{subject to} \quad v \in \mathbf{C}^n\\
&\qquad\qquad \gamma_i^{\min} - d_i \le v^* S_i v \le \gamma_i^{\max} - d_i, \quad i \in \mathcal{V},\\
&\qquad\qquad v_i^{\min} \le |v_i| \le v_i^{\max}, \quad i \in \mathcal{V},\\
&\qquad\qquad -\ell_{ij}^{\max} \le v^* P_{ij} v \le \ell_{ij}^{\max}, \quad (i,j) \in \mathcal{E}.
\end{aligned}$$

It is important to note that, despite being deterministic, the above AC-OPF problem is nonconvex and NP-hard, in general. There are, however, many off-the-shelf optimization routines (e.g., Matpower [96]) that are reliable in their ability to obtain *feasible* solutions to problem (5.12).

The following Proposition establishes that a feasible solution to (5.12) can be used to construct a matrix V_0, which is guaranteed to generate a nonempty feasible region for the convex inner approximation $\mathcal{P}_{\text{III}}(V_0)$ of the RAC-OPF problem.

Proposition 5.4.4. *Let $v_0 \in \mathbf{C}^n$ be a feasible solution to* (5.12)*, and define a matrix $V_0 := v_0 e_1^*$. Then, the optimization problem $\mathcal{P}_{\text{III}}(V_0)$ has a nonempty feasible region.*

Proof. Given a feasible solution v_0 to problem (5.12), we show that there exists $x_0 \in \mathbf{R}^{2n}$ and $\Lambda_0 \in \mathbf{R}^{m \times \ell}$ such that the point (x_0, V_0, Λ_0) is feasible for problem $\mathcal{P}_{\text{III}}(V_0)$, where $V_0 = v_0 e_1^*$. Let $x_0 \in \mathbf{R}^{2n}$ be a vector given by

$$[x_0]_i := \begin{cases} \text{Re}\{v_0^* S_i v_0 + d_i\}, & \text{if } 1 \le i \le n, \\ \text{Im}\{v_0^* S_{i-n} v_0 + d_{i-n}\}, & \text{if } n < i \le 2n, \end{cases}$$

where $S_i \in \mathbf{C}^{n\times n}$ is defined in (A.4). Since v_0 is a feasible solution to (5.12), it is easy to verify that $Ex_0 \le f$, where $E \in \mathbf{R}^{4n\times 2n}$ and $f \in \mathbf{R}^{4n}$ are defined in Appendix A.3. Therefore, it suffices to show that there exists $\Lambda_0 \le 0$ such that

$$H_i(V_0, V_0) - C_i + (b_i^* x_0) e_1 e_1^* - \sum_{j=1}^{\ell} [\Lambda_0]_{ij} W_j \preceq 0, \tag{5.13}$$

for all $i \in [m]$. Since $V_0 = v_0 e_1^*$, we have

$$H_i(V_0, V_0) = V_0^* A_i V_0 = (v_0^* A_i v_0) e_1 e_1^*, \tag{5.14}$$

for all $i \in [m]$. Therefore, (5.13) takes the form

$$(v_0^* A_i v_0 + b_i^* x_0) e_1 e_1^* - C_i - \sum_{j=1}^{\ell} [\Lambda_0]_{ij} W_j \preceq 0, \tag{5.15}$$

for all $i \in [m]$. Fix an $i \in [m]$ and consider the constraint

$$v^* A_i v + b_i^* x \le c_i^* \xi, \quad \forall\, \xi \in \Xi,$$

in problem $\mathcal{P}$. This constraint can be equivalently expressed as $v^* A_i v + b_i^* x \le \pi_i$, where π_i is the optimal value of the following convex quadratically constrained quadratic program

$$\begin{aligned} \text{minimize} \quad & \xi^* C_i \xi \\ \text{subject to} \quad & \xi \in \mathbf{R}^k \\ & \xi^* W_j \xi \ge 0, \qquad j \in [\ell], \\ & \xi_1 = 1. \end{aligned}$$

In the above optimization problem, we have used the relation $\xi_1 = 1$ to rewrite $c_i^* \xi$ as $\xi^* C_i \xi$, where $C_i \in \mathbf{R}^{k\times k}$ is defined in problem $\mathcal{P}_{\text{II}}$. By Asumption 5.2.1, there exists $\xi \in \Xi$ such that $\xi^* W_j \xi > 0$ for all $j \in [\ell]$. This is a Slater condition, which guarantees strong duality to hold between the above optimization problem and its dual. Therefore, $\pi_i = \tau_i$, where τ_i is the optimal value of the dual problem, which is given by

$$
\begin{aligned}
\text{maximize} \quad & \rho_i + \gamma_i \\
\text{subject to} \quad & \rho_i \in \mathbf{R},\ \gamma_i \in \mathbf{R},\ [\Lambda_0]_{ij} \in \mathbf{R},\ j \in [\ell] \\
& \begin{bmatrix} C_i + \sum_{j=1}^{\ell} [\Lambda_0]_{ij} W_j & (\rho_i/2)e_1 \\ \\ (\rho_i/2)e_1^* & -\gamma_i \end{bmatrix} \succeq 0, \\
& [\Lambda_0]_{ij} \leq 0, \quad j \in [\ell].
\end{aligned}
$$

Now, recall problem (5.12), which entails the computation of a day-ahead dispatch, which is guaranteed to be feasible for the RAC-OPF problem $\mathcal{P}$ without requiring adjustment (recourse) in real-time. Since v_0 is a feasible solution to (5.12), we must have

$$
v_0 A_i v_0 + b_i^* x_0 \leq \pi_i = \tau_i.
$$

Therefore, there exists $\rho_i, \gamma_i, [\Lambda_0]_{i1}, \ldots, [\Lambda_0]_{i\ell} \in \mathbf{R}$, such that $[\Lambda_0]_{ij} \leq 0$, for all $j \in [\ell]$, $v_0^* A_i v_0 + b_i^* x_0 \leq \rho_i + \gamma_i$, and

$$
\begin{bmatrix} C_i + \sum_{j=1}^{\ell} [\Lambda_0]_{ij} W_j & (\rho_i/2)e_1 \\ \\ (\rho_i/2)e_1^* & -\gamma_i \end{bmatrix} \succeq 0. \tag{5.16}
$$

Since $v_0^* A_i v_0 + b_i^* x_0 \leq \rho_i + \gamma_i$, it follows readily that

$$
(v_0^* A_i v_0 + b_i^* x_0) e_1 e_1^* \ \preceq \ (\rho_i + \gamma_i) e_1 e_1^*. \tag{5.17}
$$

Subtracting $C_i + \sum_{j=1}^{\ell} [\Lambda_0]_{ij} W_j$ from both sides of (5.17), we observe that its left-hand side becomes equal to (5.15). Therefore, it suffices to show that

$$
(\rho_i + \gamma_i) e_1 e_1^* - C_i - \sum_{j=1}^{\ell} [\Lambda_0]_{ij} W_j \ \preceq \ 0, \tag{5.18}
$$

as this implies that (5.15) holds. Using Lemma 3.1 in [38], the positive semidefiniteness constraint (5.16) can be equivalently described by the following set of conditions:

$$\begin{cases} C_i + \sum_{j=1}^{\ell} [\Lambda_0]_{ij} W_j \succeq 0, \\ \gamma_i \leq 0, \\ \gamma_i \Big(C_i + \sum_{j=1}^{\ell} [\Lambda_0]_{ij} W_j \Big) + (\rho_i^2/4) e_1 e_1^* \preceq 0. \end{cases} \tag{5.19}$$

Consider the case where $\gamma_i = 0$. Then, (5.16) implies that $\rho_i = 0$ since for any positive semidefinite matrix X, if $[X]_{mm} = 0$, then $[X]_{ml} = [X]_{lm} = 0$, for all $l \in [n]$ (see, for example, 7.1.10 in [31]). In this case, (5.18) coincides with the first condition in (5.19) and we are done.

Next, consider the case where $\gamma_i < 0$. By rearranging terms, the third condition in (5.19) holds if and only if

$$-C_i - \sum_{j=1}^{\ell} [\Lambda_0]_{ij} W_j \ \preceq \ \frac{\rho_i^2}{4\gamma_i} e_1 e_1^*.$$

Using the above matrix inequality, we can bound from above (in the positive semidefinite sense) the left hand side of (5.18). Thus, it suffices to show that $(\rho_i + \gamma_i + \rho_i^2/(4\gamma_i)) e_1 e_1^* \preceq 0$ or, stated equivalently,

$$\rho_i + \gamma_i + \frac{\rho_i^2}{4\gamma_i} \leq 0. \tag{5.20}$$

For a matrix $X \in \mathbf{H}^n$, let $\eta(X) \in \mathbf{R}^n$ be the vector of eigenvalues of X arranged in nonincreasing order. Since for any $X, Y \in \mathbf{H}^n$, if $X \succeq Y$, then $\eta(X) \geq \eta(Y)$ (see, for example, Corollary 7.7.4 in [31]), the third condition in (5.19) implies that

$$\rho_i^2 \leq -4\gamma_i \eta_{\max},$$

where $\eta_{\max}$ is the largest eigenvalue of $C_i + \sum_{j=1}^{\ell}[\Lambda_0]_{ij}W_j$. As this is a positive semidefinite matrix, we must also have that $\eta_{\max}$ is nonnegative. Therefore, $\rho_i \leq 2\sqrt{-\gamma_i\eta_{\max}}$. Using this inequality, we can bound the left hand-side of (5.20) from above. And therefore, it suffices to show that

$$2\sqrt{-\gamma_i\eta_{\max}} + \gamma_i - \eta_{\max} \ \leq \ 0, \tag{5.21}$$

as this implies that (5.20) holds. Indeed, consider the function $f : \mathbf{R}_+ \times \mathbf{R}_{++} \to \mathbf{R}$ given by $f(x, y) = -y + 2\sqrt{yx} - x$, where $\mathbf{R}_+,\ \mathbf{R}_{++}$ denote the sets of nonnegative and positive real numbers, respectively. To complete the proof notice that the maximum of f on $\mathbf{R}_+ \times \mathbf{R}_{++}$ occurs at $x = y$ and takes the value zero. ■

5.4.4 Recursive Convex Inner Approximations

In what follows, we describe a recursive method that builds upon our previous development to generate a sequence of cost-improving convex inner approximations to the RAC-OPF problem. Given a feasible solution $v_0 \in \mathbf{C}^n$ to the deterministic AC-OPF problem (5.12), define the matrix $V_0 \coloneqq v_0 e_1^*$, and consider a recursion of the form

$$(x_{t+1}, V_{t+1}, \Lambda_{t+1}) \ \in \ \underset{(x,V,\Lambda)\in\mathcal{F}(V_t)}{\text{argmin}} \ \text{tr}(MH_0(V, V_t)). \tag{5.22}$$

Here, $\mathcal{F}(V_t)$ denotes the feasible set of problem (5.12) parameterized by the matrix V_t. The recursive algorithm (5.22) can be interpreted as implementing a successive convex majorization-minimization method.

We have the following Proposition, which establishes two important properties of the recursive method: (i) it is guaranteed to yield a nonempty convex inner approximation to the RAC-OPF problem at each step in the recursion, and (ii) it is guaranteed to generate a sequence of *feasible* dispatch policies for the RAC-OPF problem with nonincreasing costs.

Proposition 5.4.5. *Let* $\{x_t, V_t, \Lambda_t\}_{t=0}^{\infty}$ *denote the sequence of solutions generated by the recursion in* (5.22). *The following properties hold for each step* t *of the recursion.*

(i) Nonemptiness: $\mathcal{F}(V_t) \neq \emptyset$.

(ii) Cost montonicity: $\text{tr}(MV_t^* A_0 V_t) \leq \text{tr}(MV_{t-1}^* A_0 V_{t-1})$.

Proof. For notational brevity, let us first define matrices

$$F_i(Z, x, V, \Lambda) := H_i(V, Z) - C_i + (b_i^* x) e_1 e_1^* - \sum_{j=1}^{\ell} [\Lambda]_{ij} W_j,$$

for each $i \in [m]$. In addition, we let

$$J(V, V_t) := \text{tr}(MH_0(V, V_t)).$$

The proof is by induction on t.

Base of induction: Consider the case $t = 0$. By Propostion 5.4.4, there exists $x_0 \in \mathbf{R}^{2n}$ and $\Lambda_0 \in \mathbf{R}^{m \times \ell}$ such that $(x_0, V_0, \Lambda_0) \in \mathcal{F}(V_0)$. Hence $\mathcal{F}(V_0) \neq \emptyset$ and we are done with (i) for the base of induction. Let (x_1, V_1, Λ_1) be a solution to problem (5.22). By its optimality, we must have $J(V_1, V_0) \leq J(V, V_0)$ for all $(u, V, \Lambda) \in \mathcal{F}(V_0)$. Since $(x_0, V_0, \Lambda_0) \in \mathcal{F}(V_0)$, we obtain in particular

$$J(V_1, V_0) \leq J(V_0, V_0) = \text{tr}(MV_0^* A_0 V_0).$$

It remains to show that $\text{tr}(MV_1^* A_0 V_1) \leq J(V_1, V_0)$ in order to establish (ii) for the base of induction. Indeed, letting $V = V_1$ and $Z = V_0$ in Lemma 5.4.2(ii), we readily obtain the desired inequality.

Step of Induction: Let $t = s$ and suppose that $\mathcal{F}(V_j) \neq \emptyset$ for all $j < s$. We must show that $\mathcal{F}(V_s) \neq \emptyset$ in order to establish (i). To do so, we use the assumption that $\mathcal{F}(V_{s-1}) \neq \emptyset$, together with Lemma 5.4.2(i) to show that $(x_s, V_s, \Lambda_s) \in \mathcal{F}(V_s)$. Let $(x_s, V_s, \Lambda_s) \in \mathcal{F}(V_{s-1})$ be a solution to (5.22). This point is guaranteed to exist since $\mathcal{F}(V_{s-1}) \neq \emptyset$ by assumption. Thus, $Ex_s \leq f$, $\Lambda_s \leq 0$, and

$$F_i(V_{s-1}, x_s, V_s, \Lambda_s) \preceq 0, \quad \forall\, i \in [m]. \tag{5.23}$$

Fix an $i \in [m]$. By using Lemma 5.4.2((i)) at $V = V_s$ and $Z = V_{s-1}$, we obtain

$$V_s^* A_i V_s \ \preceq \ H_i(V_s, V_{s-1}). \tag{5.24}$$

Adding $-C_i + (b_i^* x_s) e_1 e_1^* - \sum_{j=1}^{\ell} [\Lambda_s]_{ij} W_j$ to both sides of (5.24), yields

$$V_s^* A_i V_s - C_i + (b_i^* x_s) e_1 e_1^* - \sum_{j=1}^{\ell} [\Lambda_s]_{ij} W_j \preceq F_i(V_{s-1}, x_s, V_s, \Lambda_s) \preceq 0. \tag{5.25}$$

where the last relation follows from (5.23). Consider now the feasible set $\mathcal{F}(V_s)$. We claim that $(x_s, V_s, \Lambda_s) \in \mathcal{F}(V_s)$. Indeed $Ex_s \leq f$, $\Lambda_s \leq 0$ and $F_i(V_s, x_s, V_s, \Lambda_s) \preceq 0$ for all $i \in [m]$, where the last relation follows from (5.25) and the fact that $H_i(V_s, V_s) = V_s^* A_i V_s$. This completes the proof of the step of induction for part (i).

We will now prove the step of induction for part (ii). We must show that $\mathrm{tr}(MV_s^* A_0 V_s) \leq \mathrm{tr}(MV_{s-1}^* A_0 V_{s-1})$ for any solution (x_s, V_s, Λ_s) of (5.22). Let (x_s, V_s, Λ_s) be one such solution. Since it is optimal, we must have

$$J(V_s, V_{s-1}) \leq J(V, V_{s-1}), \quad \forall\, (x, V, \Lambda) \in \mathcal{F}(V_{s-1}).$$

In the step of induction of part (i), we have shown that $(x_{s-1}, V_{s-1}, \Lambda_{s-1}) \in \mathcal{F}(V_{s-1})$. Therefore, we obtain in particular

$$J(V_s, V_{s-1}) \leq J(V_{s-1}, V_{s-1}) = \mathrm{tr}(MV_{s-1}^* A_0 V_{s-1}).$$

It remains to show that $\mathrm{tr}(MV_s^* A_0 V_s) \leq J(V_s, V_{s-1})$. This follows readily by setting $V = V_s$ and $Z = V_{s-1}$ in Lemma 5.4.2(ii). ∎

5.5 Convex Outer Approximation

In this section, we develop a method for constructing finite-dimensional convex relaxations (outer approximations) to the RAC-OPF problem. These relaxations enable the computation of lower bounds on the optimal value of the RAC-OPF problem. Such lower bounds can, in turn, be used to bound the suboptimality incurred by the feasible affine policies proposed in Section 5.4.2. In particular, if the gap between the optimal values of the outer and inner approximations is small, we have an a posteriori certificate of near optimality of the feasible solution computed.

5.5.1 Second Order Cone Outer Approximations

In this Section, we develop a finite-dimensional second-order conic relaxation to the two-stage RAC-OPF problem formulated in Section 5.3. As the initial step in the derivation of a tractable relaxation to the RAC-OPF problem, we reformulate program $\mathcal{P}$ as a two-stage robust rank one constrained semidefinite program of the following form:

$$\begin{aligned}
\text{minimize} \quad & \mathbb{E}\Big[\text{tr}\big(A_0 V(\boldsymbol{\xi})\big)\Big] && (5.26)\\
\text{subject to} \quad & x \in \mathbf{R}^{2n},\ V \in \mathbf{L}^2_{k,n\times n}, \\
& Ex \le f, \\
& \left.\begin{aligned}
& \text{tr}\big(A_i V(\boldsymbol{\xi})\big) + b_i^* x \le c_i^* \boldsymbol{\xi}, \quad i = 1, \dots, m \\
& V(\boldsymbol{\xi}) \in \mathbf{H}^n_+, \\
& \text{rank}\big(V(\boldsymbol{\xi})\big) \le 1,
\end{aligned}\right\} \mathbb{P}\text{-a.s..}
\end{aligned}$$

This reformulation arises by using the invariance of trace under cyclic permutations to massage each quadratic function to obtain $v(\boldsymbol{\xi})^* A_i v(\boldsymbol{\xi}) = \text{tr}(A_i v(\boldsymbol{\xi}) v(\boldsymbol{\xi})^*)$. By performing a change of variables $V(\boldsymbol{\xi}) := v(\boldsymbol{\xi}) v(\boldsymbol{\xi})^*$, we obtain the robust rank one constrained semidefinite program (5.26).

The non-convexity in problem (5.26) is concentrated in the rank constraint and the relaxation to a robust semidefinite program entails the removal of the rank one constraint. Robust semidefinite programs, however, are computationally intractable, in general [6]. The difficulty derives from the need to verify the nonnegativity of a nonlinear concave function over the convex uncertainty set. Therefore, we take the approach of replacing the positive semidefinite cone with a polyhedral cone, which contains the positive semidefinite cone. That is to say, we require that

$$V(\boldsymbol{\xi}) \in \mathbf{P}, \ \mathbb{P}\text{-a.s.},$$

where $\mathbf{P} \supseteq \mathbf{H}^n_+$ is a polyhedral cone. It is defined as

$$\mathbf{P} := \bigcap_{i=1}^{p} \left\{ W \in \mathbf{H}^n \mid \text{tr}(Z_i W) \geq 0 \right\}, \tag{5.27}$$

where $Z_i \in \mathbf{H}^n_+$ for all $i = 1, \ldots, p$. Henceforth, we refer to $\mathbf{P}$ as the *outer polyhedral cone*. Since the positive semidefinite cone is a self-dual cone, it is straightforward to see that $\mathbf{P} \supseteq \mathbf{H}^n_+$. One example of an outer polyhedral cone is the cone of Hermitian matrices with positive diagonal entries. It is obtained by setting $Z_i = e_i e_i^*$ for all $i = 1, \ldots, n$.

The above relaxation of the positive semidefiniteness constraint yields the following robust linear program, whose optimal value stands as a lower bound to the optimal value of the

RAC-OPF problem.

$$
\begin{aligned}
\text{minimize} \quad & \mathbb{E}\Big[\mathrm{tr}\big(A_0 V(\boldsymbol{\xi})\big)\Big] && (5.28)\\
\text{subject to} \quad & x \in \mathbf{R}^{2n},\ V \in \mathbf{L}^2_{k,n\times n} \\
& Ex \leq f, \\
& \left.\begin{aligned} & \mathrm{tr}\big(A_i V(\boldsymbol{\xi})\big) + b_i^* x \leq c_i^* \boldsymbol{\xi}, \quad i = 1, \ldots, m \\ & V(\boldsymbol{\xi}) \in \mathbf{P}, \end{aligned}\right\} \ \mathbb{P}\text{-a.s.}
\end{aligned}
$$

The robust linear program (5.28), remains to be intractable as it involves infinite-dimensional decision variables and almost-sure constraints. The finite-dimensional conic relaxation of the RAC-OPF probelm is obtained from program (5.28) through two key approximation steps. The first entails restricting the space of dual policies corresponding to the almost-sure constraints in problem (5.28) to be linear in the random vector $\boldsymbol{\xi}$. The second amounts to relaxing a set of moment feasibility constraints, which arise by the restriction to dual linear policies. These relaxation steps are developed in the proof of Theorem 5.5.1, which follows largely from arguments in [40]. Before stating Theorem 5.5.1, it will be convenient to define the cone generated by the uncertainty set Ξ. It is given by

$$\mathrm{cone}(\Xi) = \{z \in \mathbf{R}^k \mid z_1 \geq 0,\ z^\top W_j z \geq 0,\ \forall\, j = 1, \ldots, \ell\}.$$

We have the following Theorem, which provides a finite-dimensional conic relaxation for the RAC-OPF probelm.

Theorem 5.5.1. *The optimal value of the second-order cone program*

$$
\begin{array}{lll}
\text{minimize} & e_1^* M \displaystyle\sum_{t=1}^{k} \text{tr}(A_0 V_t) e_t & \\
\text{subject to} & x \in \mathbf{R}^{2n},\ S \in \mathbf{R}^{m\times k},\ \{V_i\}_{i=1}^{k} \in \mathbf{H}^n & \\
& Ex \leq f, & \\
& S^* e_i + \displaystyle\sum_{t=1}^{k} \text{tr}(A_i V_t) e_t + b_i^* x e_1 = c_i, & i = 1, \ldots, m, \\
& M S^* e_i \in \text{cone}(\Xi), & i = 1, \ldots, m, \\
& M \displaystyle\sum_{t=1}^{k} \text{tr}(Z_i V_t) e_t \in \text{cone}(\Xi), & i = 1, \ldots, p.
\end{array}
\tag{5.29}
$$

is a lower bound on the optimal value of RAC-OPF.

Proof. Starting from the two-stage robust linear program 5.28, we apply two main relaxation steps which yield the finite-dimensional second order cone program 5.29. The first amounts to restricting the space of dual policies corresponding to the almost-sure constraint in problem (5.28) to be linear in the random vector $\boldsymbol{\xi}$. The second amounts to relaxing a set of moment feasibility constraints.

Before we delve into the details of the proof, it will first be convenient to define matrices $A_{m+1}, \ldots, A_{m+p} \in \mathbf{H}^n$ and vectors $b_{m+1}, \ldots, b_{m+p} \in \mathbf{R}^{2n}$ and $c_{m+1}, \ldots, c_{m+p} \in \mathbf{R}^k$ as follows

$$
A_{m+i} \coloneqq -Z_i, \qquad b_{m+i} \coloneqq 0, \qquad c_{m+i} \coloneqq 0, \quad i = 1, \ldots, p. \tag{5.30}
$$

Using these definitions, we can express the constraint $V(\boldsymbol{\xi}) \in \mathbf{P}^n,\ \mathbb{P}\text{-a.s.}$ in (5.28) follows:

$$
\text{tr}\big(A_{m+i} V(\boldsymbol{\xi})\big) + b_{m+i}^* x \leq c_{m+i}^* \boldsymbol{\xi}, \quad i = 1, \ldots, p,\ \mathbb{P}\text{-a.s.},
$$

Therefore, problem (5.28) can be stated as follows:

$$
\begin{aligned}
&\text{minimize} \quad \mathbb{E}\Big[\mathrm{tr}\big(A_0 V(\boldsymbol{\xi})\big)\Big] \\
&\text{subject to} \quad x \in \mathbf{R}^{2n},\ V \in \mathbf{L}^2_{k,n\times n}, s \in \mathcal{L}^2_{k,m+p} \\
&\qquad\qquad\quad Ex \le f, \\
&\qquad\qquad\quad \left.\begin{aligned} &s_i(\boldsymbol{\xi}) + \mathrm{tr}\big(A_i V(\boldsymbol{\xi})\big) - (b_i^* x e_1 - c_i)^* \boldsymbol{\xi} = 0, \quad i = 1, \ldots, m+p \\ &s(\boldsymbol{\xi}) \ge 0 \end{aligned}\right\} \mathbb{P}\text{-a.s.},
\end{aligned}
\tag{5.31}
$$

where we have introduced slack variables $s_1, \ldots, s_{m+p} \in \mathcal{L}^2_{k,m+p}$, to express the inequality constraints as equality constraints. We now argue that a lower bound to the optimal value of problem (5.31) is obtained by restricting the dual multiplier functions corresponding to the equality constraints in (5.31) to be linear in ξ. To do so, it is necessary to first dualize these equality constraints to obtain the following equivalent min-max reformulation of (5.31), where the inner maximization is over the dual policies $\mathsf{y}_i \in \mathcal{L}^2_{k,1},\ i \in [m']$

$$
\begin{aligned}
&\text{minimize} \quad \mathbb{E}\Big[\mathrm{tr}\big(A_0 V(\boldsymbol{\xi})\big)\Big] + \sum_{i=1}^{m+p} \sup_{\mathsf{y}_i \in \mathcal{L}^2_{k,1}} \mathbb{E}\Big[\mathsf{y}_i(\boldsymbol{\xi})\left(s_i(\boldsymbol{\xi}) + \mathrm{tr}(A_i V(\boldsymbol{\xi})) - (b_i^* x e_1 - c_i)^* \boldsymbol{\xi}\right)\Big] \\
&\text{subject to} \quad V \in \mathbf{L}^2_{k,n\times n},\ s \in \mathcal{L}^2_{k,m+p} \\
&\qquad\qquad\quad s(\boldsymbol{\xi}) \ge 0, \ \ \mathbb{P}\text{-a.s.}
\end{aligned}
\tag{5.32}
$$

We then restrict the functional form of the dual policies $\mathsf{y}_i(\boldsymbol{\xi}),\ i = 1, \ldots, m+p$ to be linear in $\boldsymbol{\xi}$. That is to say, we require that

$$
\mathsf{y}_i(\boldsymbol{\xi}) = y_i^* \boldsymbol{\xi},
$$

for all $i = 1, \ldots, m+p$, where $y_i \in \mathbf{R}^k$. This restriction to dual affine policies implies that

$$
\begin{aligned}
&\sup_{\mathsf{y}_i \in \mathcal{L}^2_{k,1}} \mathbb{E}\Big[\mathsf{y}_i(\boldsymbol{\xi})\big(s_i(\boldsymbol{\xi}) + \mathrm{tr}(A_i V(\boldsymbol{\xi})) - (b_i^* x e_1 - c_i)^* \boldsymbol{\xi}\big)\Big] \\
&\qquad\qquad \ge \sup_{y_i \in \mathbf{R}^k} y_i^* \mathbb{E}\Big[\big(s_i(\boldsymbol{\xi}) + \mathrm{tr}\big(A_i V(\boldsymbol{\xi})\big) - (b_i^* x e_1 - c_i)^* \boldsymbol{\xi}\big)\boldsymbol{\xi}\Big],
\end{aligned}
\tag{5.33}
$$

for all $i = 1, \ldots, m+p$. By replacing each supremum term in the objective function of (5.32) with the corresponding supremum term in the right hand side of (5.33) we obtain an optimization problem whose optimal value is a lower bound to the optimal value of (5.32). The maximization on the right-hand side of (5.33) admits a closed-form solution. Namely, it is equal to zero if

$$\mathbb{E}\left[\left(\operatorname{tr}\big(A_i V(\boldsymbol{\xi})\big) - (b_i^* x e_1 - c_i)^* \boldsymbol{\xi} + s_i(\boldsymbol{\xi})\right)\boldsymbol{\xi}\right] = 0,$$

and equal to plus infinity, otherwise. Therefore, the optimal value of problem (5.32) is bounded from below by the optimal value of the following optimization problem

$$\begin{aligned}
&\text{minimize} && \mathbb{E}\Big[\operatorname{tr}\big(A_0 V(\boldsymbol{\xi})\big)\Big] \\
&\text{subject to} && x \in \mathbf{R}^{2n},\ V \in \mathcal{L}^2_{k,n\times n},\ s \in \mathcal{L}^2_{k,m+p}, \\
& && \mathbb{E}\Big[\Big(\operatorname{tr}\big(A_i V(\boldsymbol{\xi})\big) - (b_i^* x e_1 - c_i)^* \boldsymbol{\xi} + s_i(\boldsymbol{\xi})\Big)\boldsymbol{\xi}\Big] = 0, \quad i = 1, \ldots, m+p, \\
& && s(\boldsymbol{\xi}) \geq 0, \qquad \mathbb{P}\text{-a.s.}.
\end{aligned} \tag{5.34}$$

Problem (5.34) is described by finitely many equality constraints containing an expectation term. The next step entails reformulating the above optimization problem to eliminate these expectation terms. To do so, we introduce variables $S \in \mathbf{R}^{(m+p)\times k}$ and $\{V_i\}_{i=1}^k \in \mathbf{H}^n$, which are determined by $s(\boldsymbol{\xi})$ and $V(\boldsymbol{\xi})$, respectively, through the following expressions:

(i) $MS^* e_i = \mathbb{E}[s_i(\boldsymbol{\xi})\boldsymbol{\xi}]$, for all $i = 1, \ldots, m+p$

(ii) $\displaystyle\sum_{t=1}^{k} (e_j^* M e_t) V_t = \mathbb{E}\big[\boldsymbol{\xi}_j V(\boldsymbol{\xi})\big]$, for all $j = 1, \ldots, k$.

where recall that $M = \mathbb{E}[\boldsymbol{\xi}\boldsymbol{\xi}^*]$ is the second-order moment matrix. Using definitions (i) and (ii), the linearity of the trace operator, and the fact that $\boldsymbol{\xi}_1 = 1$, $\mathbb{P}$-a.s., the objective function of problem (5.34) can be expressed as follows:

$$\mathbb{E}\Big[\text{tr}\big(A_0 V(\boldsymbol{\xi})\big)\Big] = \text{tr}\Big(A_0 \mathbb{E}\big[V(\boldsymbol{\xi})\boldsymbol{\xi}_1\big]\Big) = e_1^* M \sum_{t=1}^{k} e_t \text{tr}(A_0 V_t) \tag{5.35}$$

Similar arguments allow us to reformulate the equality constraints in (5.34) as follows:

$$S^* e_i + \sum_{t=1}^{k} \text{tr}(A_i V_t) e_t + b_i^* x = c_i, \quad i = 1, \ldots, m+p, \tag{5.36}$$

where we have used the fact that M is positive definite and therefore invertible. Recall now the definition of the matrices A_{m+i} and the vectors c_{m+i} and b_{m+i} for $i = 1, \ldots, p$. It follows from (5.36) that

$$S^* e_{m+i} = \sum_{t=1}^{k} \text{tr}(Z_i V_t) e_t, \quad i = 1, \ldots, p. \tag{5.37}$$

Let us now define the following convex cone in $\mathbf{R}^k$

$$\mathcal{K} := \{z \in \mathbf{R}^k \mid \exists\, s \in \mathcal{L}_{k,1}^2 \text{ s.t. } z = \mathbb{E}[s(\boldsymbol{\xi})\boldsymbol{\xi}],\ s(\boldsymbol{\xi}) \geq 0,\ \mathbb{P}\text{-a.s.}\}.$$

We will use this cone to reformulate problem (5.34) as a conic optimization problem over the cone $\mathcal{K}$. To do so, first observe that $MS^* e_i = \mathbb{E}[s_i(\boldsymbol{\xi})\boldsymbol{\xi}]$, $s_i(\boldsymbol{\xi}) \geq 0$ for all $i \in 1, \ldots, m+p$ if and only if

(i) $MS^* e_i \in \mathcal{K}$, for all $i = 1, \ldots, m$, and

(ii) $M \sum_{t=1}^{k} \text{tr}(Z_i V_t) e_t \in \mathcal{K}$, for all $i = 1, \ldots, p$,

where in (ii) we have used equation (5.37). Using (5.35), (5.36), and (i)-(ii) above, we obtain the following reformulation for problem (5.34) as a conic optimization problem over $\mathcal{K}$.

$$
\begin{aligned}
\text{minimize} \quad & e_1^* M \sum_{t=1}^{k} \operatorname{tr}(A_0 V_t) e_t && (5.38)\\
\text{subject to} \quad & x \in \mathbf{R}^{2n},\ S \in \mathbf{R}^{m\times k},\ \{V_t\}_{t=1}^k \in \mathbf{H}^n,\ V \in \mathcal{L}^2_{k,n\times n} \\
& S^* e_i + \sum_{t=1}^{k} \operatorname{tr}(A_i V_t) e_t + b_i^* x = c_i, && i = 1, \dots, m,\\
& M S^* e_i \in \mathcal{K}, && i = 1, \dots, m,\\
& M \sum_{t=1}^{k} \operatorname{tr}(Z_i V_t) e_t \in \mathcal{K}, && i = 1, \dots, p,\\
& \sum_{t=1}^{k} (e_j^* M e_t) V_t = \mathbb{E}[\boldsymbol{\xi}_j V(\boldsymbol{\xi})], && j = 1, \dots, k.
\end{aligned}
$$

We now argue that both the last constraint and the policy $V \in \mathcal{L}^2_{k,n\times n}$ are redundant and can therefore be eliminated from (5.38). Indeed, given any feasible solution $\{V_t\}_{t=1}^k$ to problem (5.38), the linear function

$$V(\boldsymbol{\xi}) := \sum_{t=1}^{k} \boldsymbol{\xi}_t V_t,$$

satisfies the last constraint of problem (5.38). Problem (5.38) remains intractable as verifying existence of a vector z in $\mathcal{K}$ entails checking for the existence of a function s in an infinite-dimensional decision space which satisfies the moment constraint $z = \mathbb{E}[s(\boldsymbol{\xi})\boldsymbol{\xi}]$. The final step in the relaxation involves replacing the cone $\mathcal{K}$ by a cone containing it and for which there exist efficient algorithms for linear optimization over its affine slices. We have the following Lemma from [40].

Lemma 5.5.2. *The cone $\mathcal{K}$ satisfies the following relation:*

$$\emptyset \neq \mathit{int}\big(\operatorname{cone}(\Xi)\big) \subseteq \mathcal{K} \subseteq \operatorname{cone}(\Xi).$$

A direct application of Lemma 5.5.2 yields the desired relaxation for the RAC-OPF problem.

■

We remark that the last set of constraints in program (5) imply that outer polyhedral cones defined by a large number of halfspaces yield finite-dimensional conic relaxations having a large number of constraints.

5.5.2 Recursive Convex Outer Approximations

In the previous section, we proposed a method for constructing a finite-dimensional second-order cone relaxation to the RAC-OPF probelm. The effectiveness of this relaxation, however, depends critically on the choice of the polyhedral cone $\mathbf{P}$, and for any given problem it is unclear what the best choice for $\mathbf{P}$ is. A naive approach might entail the construction of a hierarchy of inner and outer polyhedral cones via a uniform discretization of the boundary of the cross polytope [46]. For high levels in the hierarchy, however, this approach can create computational inefficiencies due to the large number of half-spaces defining the outer polyhedral cone. In particular, such polyhedral cones yield finite-dimensional programs (5.29) with a large number of constraints.

In this section, we build upon the previous approach by exploring polyhedral approximations of the positive semidefinite cone that are adaptively guided by the objective function. More precisely, starting with a coarse outer polyhedral cone $\mathbf{P}$ (i.e., the cone of Hermitian matrices with nonnegative diagonal entries), we prescribe a recursive method, which uses the solution at the current iteration step to refine this cone. Specifically, at each iteration step, we identify a number of matrices, which are constructed from a primal-dual optimal solution pair to the finite-dimensional conic linear program, which are shown to lie on the boundary of the outer polyhedral cone. We project these matrices onto the positive semidefinite cone and refine the outer polyhedral cone by intersecting it with the half-spaces corresponding to the supporting hyperplanes at said projection points. The resulting feasible set of the finite-dimensional conic linear program over the refined outer polyhedral cone is shown to exclude the optimal solution at the previous iteration step.

As the initial step in the development of the recursive method, we consider the dual program to the second-order cone relaxation (5.29), which is given by

$$\begin{aligned}
\text{maximize} \quad & \sum_{i=1}^{m} \eta_i^* M c_i - f^* \lambda \\
\text{subject to} \quad & \{\theta_i\}_{i=1}^{p}, \in \mathbf{R}^k,\ \{\eta_i\}_{i=1}^{m} \in \mathbf{R}^k,\ \lambda \in \mathbf{R}^{4n}, \\
& E^* \lambda + \sum_{i=1}^{m} e_1^* M \eta_i b_i = 0, \\
& \sum_{i=1}^{p} \theta_i^* M e_t Z_i - \sum_{i=1}^{m} \eta_i^* M e_t A_i = e_1^* M e_t A_0, \quad t = 1, \dots, k, \\
& \eta_i \in \operatorname{cone}(\Xi)^+, \quad i = 1, \dots, m, \\
& \theta_i \in \operatorname{cone}(\Xi)^+, \quad i = 1, \dots, p, \\
& \lambda \geq 0.
\end{aligned} \tag{5.39}$$

Here, $\eta_i,\ i = 1, \dots, m$ and $\theta_i,\ i = 1, \dots, p$ are the dual multipliers of the penultimate and last constraint of problem (5.29), respectively.

Given a feasible solution $\{V_i\}_{i=1}^{k} \in \mathbf{H}^n$ to program (5.29), we define the following set of matrices.

$$\mathcal{O}(V_1, \dots, V_k) := \left\{ \sum_{i=1}^{k} (M\theta)_i V_i \;\middle|\; \theta \in \operatorname{cone}(\Xi)^+ \right\}.$$

In the following Lemma, we collect some important properties of the set $\mathcal{O}(V_1, \dots, V_k)$.

Lemma 5.5.3. *The following properties hold:*

(i) Let $\{V_i\}_{i=1}^k \in \mathbf{H}^n$ be a feasible solution the conic program (5.29). Then,

$$\mathcal{O}(V_1, \ldots, V_k) \subseteq \mathbf{P}^n.$$

In addition, $\mathcal{O}(V_1, \ldots, V_k) \subseteq \mathbf{H}_+^n$, if and only if

$$M \sum_{i=1}^{k} \mathrm{tr}(ZV_i)e_i \in \mathrm{cone}(\Xi), \quad \forall\, Z \in \mathbf{H}_+^n. \tag{5.40}$$

(ii) Let $(\{V_t^\star\}_{t=1}^k, \{\theta_j^\star\}_{j=1}^p)$ be a primal-dual optimal solution pair to programs (5.29) and (5.39). For each $j = 1, \ldots, p$, the matrix $V^j \in \mathcal{O}(V_1^\star, \ldots, V_k^\star)$, which is defined as

$$V^j := \sum_{i=1}^{k} (M\theta_j^\star)_i V_i^\star, \tag{5.41}$$

is on the hyperplane $\{X \in \mathbf{H}^n \mid \mathrm{tr}(Z_j X) = 0\}$.

Remark 13. *If the uncertainty set Ξ is polytopic, then (5.40) can be equivalently represented by a set of linear matrix inequalities. In particular, $\mathcal{O}(V_1, \ldots, V_k) \subseteq \mathbf{H}_+^n$ if and only if*

(i) $\sum_{i=1}^k e_1^* M e_i V_i \in \mathbf{H}_+^n$, and

(ii) $\displaystyle\sum_{i=1}^{k} e_j^* Q M e_i V_i \in \mathbf{H}_+^n$, for all $j = 1, \ldots, k$.

The last constraint in (5.29) and the necessary and sufficient condition (5.40) point in the direction of a recursive method for refining the outer polyhedral cone. Namely, we would like to find a supporting hyperplane to the positive semidefinite cone such that $\mathcal{O}(V_1, \ldots, V_k)$ is not contained in the corresponding half-space containing $\mathbf{H}_+^n$. We call such a hyperplane a *cutting plane* for $\mathcal{O}(V_1, \ldots, V_k)$. More precisely, a cutting plane for $\mathcal{O}(V_1, \ldots, V_k)$ is defined to be a hyperplane characterized by a matrix $Z \in \mathbf{H}_+^n$ and passing

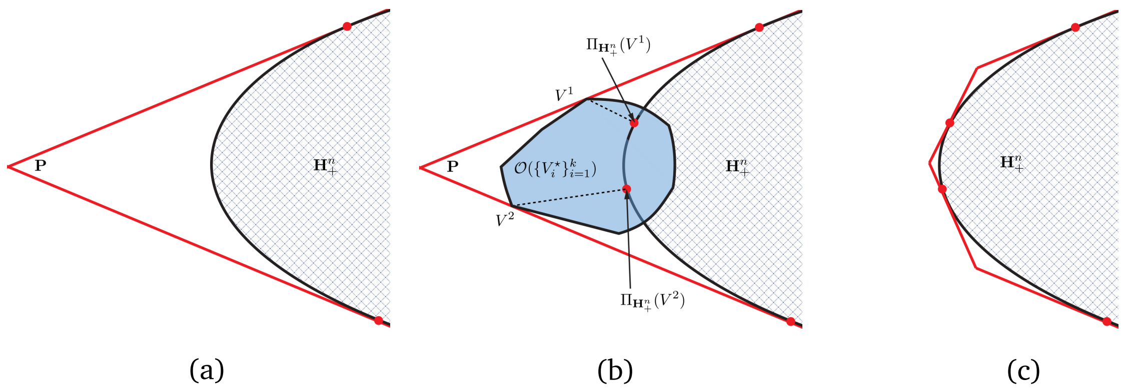


Figure 5.1: Figure 5.1(a) shows the positive semidefinite cone and the outer polyhedral cone $\mathbf{P}$, which is described by two half-spaces. In Figure 5.1(b), we visualize the set $\mathcal{O}(V_1^\star, \dots, V_k^\star))$ induced by an optimal solution $\{V_i^\star\}_{i=1}^k$ to the 5.29. We identify two matrices, V^1 and V^2 in $\mathcal{O}(V_1^\star, \dots, V_k^\star)$ that lie on the boundary of $\mathbf{P}$. The supporting hyperplanes to $\mathbf{H}_+^n$ at $\Pi_{\mathbf{H}_+^n}(V^1)$ and $\Pi_{\mathbf{H}_+^n}(V^2)$ cut off parts of $\mathcal{O}(V_1^\star, \dots, V_k^\star))$ that do not contain positive semidefinite matrices. In Figure 5.1(c), we visualize the refined outer polyhedral cone obtained by intersecting $\mathbf{P}$ with the half-spaces corresponding to the supporting hyperplanes.

through the origin, such that

$$\mathcal{O}(V_1, \dots, V_k) \not\subseteq \{X \in \mathbf{H}^n \mid \operatorname{tr}(ZX) \geq 0\}. \tag{5.42}$$

An important consequence of (5.42) is that the conic relaxation (5.29), which is obtained from approximating the positive semdefinite cone by $\mathbf{P} \cap \{X \in \mathbf{H}^n \mid \operatorname{tr}(ZX) \geq 0\}$ will not contain $\{V_i\}_{i=1}^k$ in its feasible set. We have the following Proposition.

Proposition 5.5.4. *Let $\{V_t\}_{t=1}^k$ be a feasible solution to* (5.29). *If*

$$\{X \in \mathbf{H}^n \mid \operatorname{tr}(ZX) = 0\}$$

is a cutting plane for $\mathcal{O}(V_1, \dots, V_k)$, then

$$\sum_{t=1}^k \operatorname{tr}(ZV_t) e_t \notin \operatorname{cone}(\Xi).$$

Proof. Since

$$\mathcal{O}(V_1, \dots, V_k) \not\subseteq \{X \in \mathbf{H}^n \mid \operatorname{tr}(ZX) \geq 0\},$$

it follows that there exists a $\theta \in \mathrm{cone}(\Xi)^+$ such that

$$0 > \sum_{i=1}^{k} \theta_i \mathrm{tr}(ZV_i) = \left(\sum_{i=1}^{k} \mathrm{tr}(ZV_i) e_i \right)^{\top} \theta.$$

Since $\theta \in \mathrm{cone}(\Xi)^+$, the desired exclusion follows from the definition of the dual cone. ■

Given a primal-dual optimal solution pair $(\{V_i^\star\}_{i=1}^k), \{\theta_j^\star\}_{j=1}^p)$ to programs (5.29) and (5.39), we propose a method for constructing cutting planes for $\mathcal{O}(V_1^\star, \dots, V_k^\star)$. Lemma 5.5.3(ii) identifies a set of matrices, $\{V^j\}_{j=1}^p$ in $\mathcal{O}(V_1^\star, \dots, V_k^\star)$, which lie on the boundary of the outer polyhedral cone $\mathbf{P}$. The following Proposition shows that if V^j is not positive semidefinite, then the supporting hyperplane to $\mathbf{H}_+^n$ at $\Pi_{\mathbf{H}_+^n}(V^j)$ is a cutting plane for $\mathcal{O}(V_1^\star, \dots, V_k^\star)$.

Proposition 5.5.5. *Let* $(\{V_t^\star\}_{t=1}^k, \{\theta_j^\star\}_{j=1}^p,)$ *be a primal-dual optimal solution pair to* (5.29) *and* (5.39)*. For each* $j = 1, \dots, p$*, let* $V^j \in \mathcal{O}(V_1^\star, \dots, V_k^\star)$ *be a matrix defined as in* (5.41)*. If* $V^j \notin \mathbf{H}_+^n$*, then*

$$\mathcal{O}(V_1^\star, \dots, V_k^\star) \nsubseteq \{X \in \mathbf{H}^n \mid \mathrm{tr}\big((\Pi_{\mathbf{H}_+^n}(V^j) - V^j)X\big) \geq 0\}.$$

Proof. Since $V^j \in \mathcal{O}(V_1^\star, \dots, V_k^\star)$, it suffices to show that

$$V^j \notin \{X \in \mathbf{H}^n \mid \mathrm{tr}\big((\Pi_{\mathbf{H}_+^n}(V^j) - V^j)X\big) \geq 0\},$$

as this implies that

$$\mathcal{O}(V_1^\star, \dots, V_k^\star) \nsubseteq \{X \in \mathbf{H}^n \mid \mathrm{tr}\big((\Pi_{\mathbf{H}_+^n}(V^j) - V^j)X\big) \geq 0\},$$

Let $V^j = \sum_{i=1}^n \lambda_i u_i u_i^*$ be an eigenvalue decomposition of V^j. It follows that

$$\Pi_{\mathbf{H}_+^n}(V^j) - V^j = -\sum_{i=1}^{n} \min\{0, \lambda_i\} u_i u_i^*.$$

Algorithm: Recursive Outer Approximations of RAC-OPF

Given an outer polyhedral cone $\mathbf{P}$ and a maximum number $\bar{t}$ of iterations
Initialize $t = 1$
Repeat

1. *Let* $p =$ number of half-spaces defining $\mathbf{P}$
2. *Compute.*
   - A primal-dual optimal solution pair to (5.29) and (5.39)
   - Matrices $V^1, \ldots, V^p$ defined according to (5.41)
3. *Update.*
   - $\mathbf{P} = \mathbf{P} \bigcap_{j=1}^{p} \{X \in \mathbf{H}^n \mid \mathrm{tr}\big((P_{\mathbf{H}_+^n}(V^j) - V^j)X\big) \geq 0\}$
   - $t = t + 1$

Until $V^j \in \mathbf{H}_+^n$, for all $j = 1, \ldots, p$ or $t = \bar{t}$
Output The optimal value of the finite-dimensional conic relaxation (5.29).

Table 5.1: Recursive algorithm, which yields a sequence of relaxations to the RAC-OPF problem whose optimal values is nonincreasing.

Since $V^j \notin \mathbf{H}_+^n$, it follows that

$$\begin{aligned}
0 &> \sum_{i=1}^{n} - \min\{\lambda_i, 0\}\lambda_i \\
&= \sum_{i=1}^{n} - \min\{\lambda_i, 0\} u_i^\top V^j u_i \\
&= \sum_{i=1}^{n} - \min\{\lambda_i, 0\} u_i^\top \left(\sum_{t=1}^{k} (M\theta_j)_t^\star V_t^\star \right) u_i \\
&= \mathrm{tr}\left((\Pi_{\mathbf{H}_+^n}(V^j) - V^j) \sum_{t=1}^{k} (M\theta_j)_t^\star V_t^\star \right),
\end{aligned}$$

which establishes the desired exclusion. ■

Some remarks regarding Proposition 5.5.5 are in order. First, the cutting planes defined in Proposition 5.5.5 cut off a parts of $\mathcal{O}(V_1^\star, \ldots, V_k^\star)$ that do not contain positive semidefinite matrices (see Figure 5.1b). Second, the specification of the cutting planes in Proposition 5.5.5 follows from the characterization of nearest points to nonempty closed convex sets (see, for example, Excercise 8(c) in [12]). Third, as shown in Proposition 5.5.4, *for any* index $j \in \{1, \ldots, p\}$ for which $V^j \notin \mathbf{H}^n_+$, the outer polyhedral cone

$$\mathbf{P} \cap \left\{X \in \mathbf{H}^n \;\middle|\; \mathrm{tr}\big((\Pi_{\mathbf{H}^n_+}(V^j) - V^j)X\big) \geq 0\right\}$$

yields an outer approximation to the robust semidefinite program, which does not contain $\{V_i^\star\}_{i=1}^k$ in its feasible set. Clearly, there is a computational trade-off between the number of cutting planes one chooses for refining $\mathbf{P}$ (the larger the number of cutting planes, the more the constraints of the resulting conic relaxation (5.29)) and the marginal improvement to the lower bound obtained from the resulting outer approximation. In Figure 5.1(c), we visualize the outer polyhedral cone obtained by intersecting $\mathbf{P}$ with *all* the corresponding half-spaces. Lastly, if $\{V_i^\star\}_{i=1}^k$ is uniquely optimal, then the optimal value of program (5.29) over the refined outer polyhedral cone is guaranteed to be greater than the optimal value of the corresponding approximation over $\mathbf{P}$.

In Table 5.1, we outline the steps of the proposed algorithm. Starting with a coarse outer polyhedral cone $\mathbf{P}$, the algorithm computes an optimal solution $\{V_i^\star\}_{i=1}^k$ to the finite-dimensional conic linear program (5.29). A number of cutting planes for $\mathcal{O}(V_1^\star, \ldots, V_k^\star)$ is then computed according to Proposition 5.5.5. The intersection of $\mathbf{P}$ with the corresponding half-spaces yields a refined outer polyhedral cone. The above procedure is repeated using the refined polyhedral cones until all the matrices in Proposition 5.5.5 are positive semidefinite or until a maximum number of iterations is reached.

Parameters	Units	Inflexible Generators	Intermittent Generators						Flexible Generators	
		1	2	3	4	5	6	7	8	9
α_i	\$/MW	30	0	0	0	0	0	0	50	50
$\mathrm{Re}\{g_i^{\max}\}$	MW	200	45	45	45	45	45	45	250	270
$\mathrm{Im}\{g_i^{\max}\}$	MVAR	300	†	†	†	†	†	†	300	300
$\mathrm{Re}\{g_i^{\min}\}$	MW	10	0	0	0	0	0	0	10	10
$\mathrm{Im}\{g_i^{\min}\}$	MVAR	-300	†	†	†	†	†	†	-300	-300
$\mathrm{Re}\{r_i^{\max}\} = -\mathrm{Re}\{r_i^{\min}\}$	MW	0	∞	∞	∞	∞	∞	∞	∞	∞
$\mathrm{Im}\{r_i^{\max}\} = -\mathrm{Im}\{r_i^{\min}\}$	MVAR	0	∞	∞	∞	∞	∞	∞	∞	∞
Bus index (per [71])		2	4	5	6	7	8	9	1	3

Table 5.2: Specification of each generator's location, marginal cost, and constraint parameters. The † symbol indicates that the corresponding value in the table is determined by equation (5.45).

5.6 Numerical Studies

We now illustrate the effectiveness of the proposed inner and outer approximation schemes on a nine-bus power system with varying levels of renewable resource penetration and uncertainty. We consider a modified version of the WSCC nine-bus power system. We refer the reader to [71] for its complete specification and single-line diagram. All modifications made to the original system in this chapter are summarized in Table 5.2. In particular, we assume that bus one is connected to an inflexible base-load generator, buses two through seven are connected to intermittent renewable generators, and buses eight through nine are connected to flexible peaking generators.

5.6.1 Renewable Generator Model

The real-time generating capacity of renewable generators $i \in \{2, \ldots, 7\}$ represents the only source of uncertainty in the power system being considered. Accordingly, we set

$k = 7$, and let the i^{th} element of the random vector $\boldsymbol{\xi}$ represent the *maximum active power* available to generator i in real-time. In other words,

$$\mathrm{Re}\{\underline{g}_i(\boldsymbol{\xi})\} = 0 \;\text{ and }\; \mathrm{Re}\{\overline{g}_i(\boldsymbol{\xi})\} = \boldsymbol{\xi}_i, \tag{5.43}$$

for $i = 2, \ldots, 7$. It will be convenient to our numerical analyses in the sequel to express the random vector $\boldsymbol{\xi}$ as an affine function of zero-mean random vector $\boldsymbol{\delta}$ that is uniformly distributed over a unit ball. We define this relationship according to

$$\boldsymbol{\xi} \coloneqq \mu + \sigma\boldsymbol{\delta},$$

where the random vector $\boldsymbol{\delta}$ is assumed to have support

$$\Delta \coloneqq \{\delta \in \mathbf{R}^k \mid \delta_1 = 0,\; \|\delta\|_2 \le 1\}.$$

It follows that the random vector $\boldsymbol{\xi}$ has support given by

$$\Xi = \left\{\xi \in \mathbf{R}^k \;\mid \xi - \mu \in \sigma\Delta \right\}.$$

Here, $\mu \in \mathbf{R}^k$ and $\sigma \in \mathbf{R}_+$ represent *location* and *scale parameters*, respectively. In the following study, we set $\mu_i = 15$ MW for each renewable generator $i \in \{2, \ldots, 7\}$. Qualitatively, the larger the scale parameter σ, the larger the a priori uncertainty in the real-time generating capacity of the renewable generators. The location and scale parameters are chosen in such a manner as to ensure that $\boldsymbol{\xi}_i$ respects the nameplate active power capacity limits for each renewable generator i (cf. Table 5.2). We also require that $\mu_1 = 1$ to maintain consistency with our original uncertainty model in Section 5.2.2. Finally, under the assumption that $\boldsymbol{\delta}$ has a uniform distribution, it is straightforward to show that the random vector $\boldsymbol{\xi}$ has a second-order moment matrix given by

$$M = \mu\mu^* + \left(\frac{\sigma^2}{k+1}\right)\left[\begin{array}{c|c} 0 & \\ \hline & I_{k-1} \end{array}\right].$$

Renewable energy resources, like wind and solar, employ power electronic inverters, which can produce and absorb reactive power. The limits on the maximum and minimum amount of reactive power that can be injected by a renewable generator are determined by its inverter's apparent power capacity, which we denote by $s_i^{\max} \in \mathbf{R}_+$, for each renewable generator i. It follows that the real-time complex power injection of each renewable generator i must satisfy a capacity constraint of the form

$$|g_i(\boldsymbol{\xi})| \leq s_i^{\max}. \tag{5.44}$$

As the slight oversizing of a renewable generator's apparent power rating is a standard industry practice, we set $s_i^{\max} = 1.05\mathrm{Re}\{g_i^{\max}\}$ for each renewable generator i. In order to ensure that Assumption 5.2.3 is satisfied, we enforce a more conservative form of the real-time apparent power capacity constraint (5.44) by setting the real-time reactive power limits for each renewable generator i according to

$$\mathrm{Im}\{\overline{g}_i(\boldsymbol{\xi})\} = -\mathrm{Im}\{\underline{g}_i(\boldsymbol{\xi})\} = \inf_{\xi \in \Xi} \sqrt{(s_i^{\max})^2 - \xi_i^2}. \tag{5.45}$$

The reactive power limits specified in (5.45) specify the range of reactive power injections that are guaranteed to be available to a renewable generator in real-time, regardless of the active power supplied. Finally, using the real-time active and reactive power capacity constraints specified in (5.43) and (5.45), respectively, it is straightforward to construct matrices $\underline{G}, \overline{G} \in \mathbf{C}^{n \times k}$ such that Assumption 5.2.3 is satisfied.

5.6.2 Numerical Analyses and Discussion

We begin by examining the sensitivity of the generation cost incurred under the affine recourse policies that we propose to uncertainty in renewable supply. We do so by varying

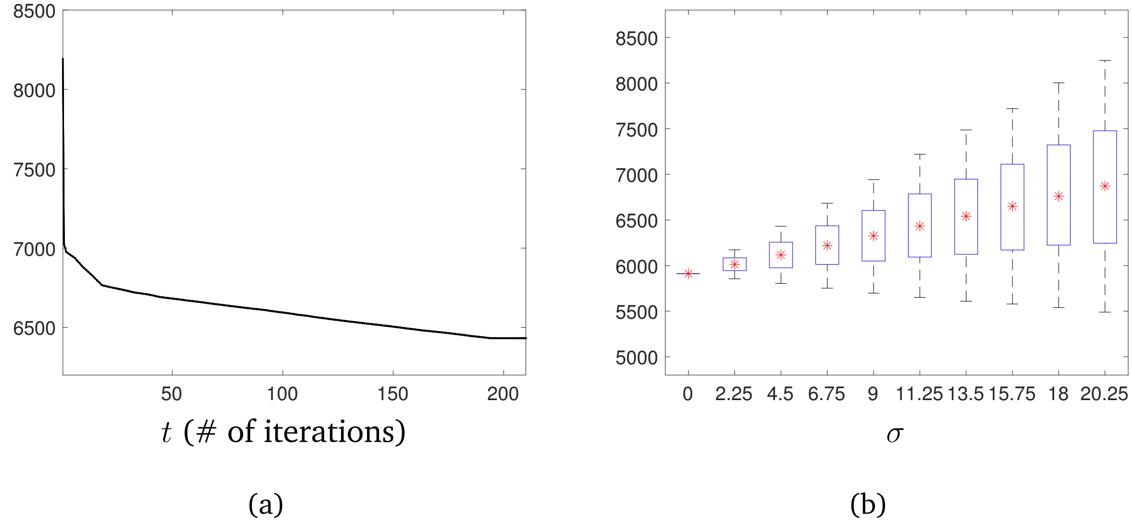
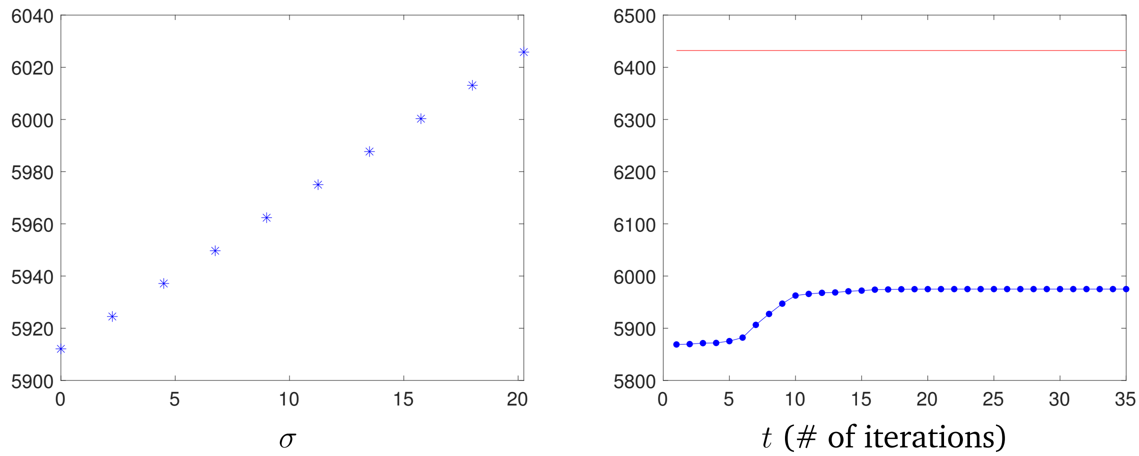


Figure 5.2: Figure (a) depicts the expected generation cost incurred by the affine recourse policy computed at each iteration of the recursive algorithm (5.22). Here, the scale parameter is set to $\sigma = 11.25$. Figure (b) depicts the expected generation cost (red star) incurred under the affine dispatch policy returned by the recursive algorithm (5.22) as a function of the scale parameter σ. For each value of σ, the figure also depicts empirical confidence intervals that are estimated from 10^4 independent realizations of the underlying random vector. The box depicts the interquartile range, while the lower and upper whiskers extend to the 5% and 95% quantiles, respectively.

the scale parameter σ from 0 to 20.25 in increments of 2.25, while keeping all other problem parameters fixed. It is worth noting that for $\sigma = 0$, there is no a priori uncertainty in the renewable supply, and the RAC-OPF problem (5.8) reduces to the deterministic AC-OPF problem (5.12). For each value of σ that we consider, we calculate an affine recourse policy according to the recursive algorithm specified in Eq. (5.22). We initialize the recursion with a feasible solution to the deterministic AC-OPF problem (5.12), which we compute using the Matpower interior point solver [96]. In Fig. 5.2(a), we plot the expected generation cost incurred by the affine recourse policy computed at each step of the inner approximation recursion (for $\sigma = 11.25$). The numerical results in Fig. 5.2(a) agree with Proposition 5.4.5, which ensures that the recursion in (5.22) will yield a sequence of feasible dispatch policies for the RAC-OPF problem with nonincreasing costs.



(a) (b)

Figure 5.3: Figure 5.3(a) depicts the optimal value of the second-order cone relaxation when the algorithm in Table 5.1 terminates as a function of the scale parameter σ. Figure 5.3(b) depicts the optimal value of the second-order cone relaxation at each iteration of the recursive algorithm in Table 5.1 (line with dots). Here, the scale parameter is set to $\sigma = 11.25$. In addition, this figure depicts the the expected generation cost incurred by the feasible affine recourse policy returned by the recursive inner approximation algorithm (5.22) for the same value of σ (solid line). By nature of these approximations, we obtain that said policy yields a cost, which is within 7.65% of the optimal value of the RAC-OPF problem.

In Fig. 5.2(b), we plot the expected generation cost (and its empirical confidence intervals) incurred by the affine dispatch policy returned by the inner approximation algorithm (5.22) versus the scale parameter σ. First, notice that the expected generation cost increases monotonically with the scale parameter. Such behavior is to be expected, as larger values of σ correspond to larger uncertainty sets Ξ. It is also worth noting the 'spread' in the cost distribution induced by the dispatch policies that we compute also increases with σ. That is to say, renewable energy resources with a large variance in their real-time generating capacity will result in a larger variance in total generating costs. Such behavior is a consequence of the risk neutrality inherent to the expected cost criterion that we treat in our formulation.

In Figure 5.3(a), we plot the optimal value of the second-order cone relaxation (5.29) returned by the recursion algorithm presented in Table 5.1 versus the scale parameter σ. These values stand as lower bounds to the optimal value of the RAC-OPF problem. We observe that the lower bounds increase monotonically with the scale parameter. This behavior is attributed to the fact that larger values of σ correspond to larger uncertainty sets Ξ.

In Figure 5.3(b), we plot (blue line with dots) the optimal value of the second-order cone relaxation at each step of the outer approximation algorithm in Table 5.1 (for $\sigma = 11.25$). In addition, we plot (solid line) the expected generation cost incurred by the affine recourse policy returned by the recursive inner approximation algorithm (5.22) for the same value of σ. First, we observe that the inner approximation recursion yields a sequence of lower bounds to the RAC-OPF problem with nondecreasing costs. Such behavior is to be expected since Proposition 5.5.5 guarantees that the optimal solution at the current iteration step is not in the feasible set of the next iteration step. In addition, Figure 5.3(b) depicts the practical value of our method. In particular, one can bound the suboptimality incurred by the feasible affine recourse policy by solving the finite-dimensional second-order cone program (5.29). In this example, the gap between the optimal values of outer and inner approximations is small. As a matter of fact, the expected cost incurred by the feasible affine recourse policy is 7.65% greater than the optimal value of the second-order cone relaxation. This gives a certificate of near optimality of the feasible solution.

5.7 Conclusions

We formulate the robust AC optimal power flow (RAC-OPF) problem as a two-stage robust optimization problem with recourse; and develop a method to approximate RAC-OPF from within by a semidefinite program. Its solution yields an affine recourse policy that is guaranteed to be feasible for RAC-OPF. We also provide an iterative optimization method that generates a sequence of feasible affine recourse policies with nonincreasing costs. In general, affine recourse policies will be suboptimal for RAC-OPF. Thus, we develop a method to approximate RAC-OPF from without by a second-order cone program. The optimal value of this optimization problem yields a lower bound to the optimal value of RAC-OPF, which can be used to bound the suboptimality incurred by the feasible affine policies proposed in this chapter. In addition, we develop a recursive method which refines the outer approximation to yield sharper lower bounds to the optimal value of RAC-OPF.

CHAPTER 6

RECURSIVE CONIC APPROXIMATIONS OF ROBUST SEMIDEFINITE PROGRAMS

6.1 Introduction

Let $\mathbf{R}^n$ be the n-dimensional Euclidean space and $\mathbf{S}^n$ ($\mathbf{S}^n_+$) space of real $n \times n$ symmetric matrices (symmetric positive semidefinite) matrices. A *robust semidefinite program* whose data is parameterized affinely in the uncertain parameter is an optimization problem of the following form

$$\begin{aligned} \underset{x \in \mathbf{R}^m}{\text{minimize}} \quad & c^\top x \\ \text{subject to} \quad & \sum_{i=1}^{k} \xi_i \mathcal{A}_i(x) \in \mathbf{S}^n_+, \quad \forall\, \xi \in \Xi, \end{aligned} \tag{RSDP}$$

where $x \in \mathbf{R}^m$ is the decision variable and $\xi \in \mathbf{R}^k$ is the uncertain parameter. The vector $c \in \mathbf{R}^m$, the affine functions $\mathcal{A}_i : \mathbf{R}^m \to \mathbf{S}^n$, and the convex *uncertainty set* $\Xi \subseteq \mathbf{R}^k$ are the given problem data. We denote by v^{opt} the optimal value of the RSDP.

The *semi-infinite* structure of the robust semidefinite program renders it computationally intractable in general. The difficulty derives from the need to verify the nonnegativity of a nonlinear concave function over the convex uncertainty set. In contrast, robust linear programs admit equivalent reformulations as finite-dimensional convex programs provided that the uncertainty set is described by an affine slice of a proper cone. For example, if the uncertainty set is polytopic, ellipsoidal, or semidefinite representable, then the robust linear program can be reformulated as a linear, second-order cone, or semidefinite program, respectively.

Although computationally intractable in general, the RSDP admits computationally tractable reformulations or inner approximations for certain characterization of the uncertainty set [6, 72]. We discuss such results in more detail in Section 6.2.2.

Contribution: We propose an approximation method, which yields an inner and outer approximation to the RSDP as a robust linear program. The approximations are obtained by approximating the semidefinite cone from within and without by a polyhedral cone. The resulting robust linear programs admit equivalent reformulations as finite-dimensional conic linear programs, provided that the uncertainty set Ξ is defined as an affine slice of a proper cone. Although possibly conservative, any solution to the robust linear program over the inner polyhedral approximation of the positive semidefinite cone will be feasible for the RSDP. Moreover, the optimal value of the robust linear program over the outer polyhedral approximation of the positive semidefinite cone serves as a lower bound on the optimal value of the RSDP. Therefore, this provides a bound on the suboptimality of the feasible point generated by the inner approximation. And if the gap between the optimal values of the outer and inner approximations is small, we have a certificate of near optimality of the feasible solution. The primary contribution of this chapter is the development of a recursive method which refines the inner and outer polyhedral cones to sharpen the approximations to the robust semidefinite program. In particular, our method is guaranteed to eliminate the optimal solution of the approximation at the current iteration step from the feasible set of the approximation at the next iteration step. And in case the inner (outer) approximation has a unique optimal solution, the cutting plane method is guaranteed to yield a sequence of decreasing (increasing) optimal values.

Organization: The remainder of this chapter is organized as follows. In Section 6.2, we present existing results pertaining to exact reformulations and approximations of robust SDPs. Sections 6.3 and 6.4 contain our main results, which include construction of inner and outer polyhedral approximations to the robust SDP and a recursive method to sharpen the approximations. In Section 6.5 we demonstrate the proposed approximations on the robust linear estimation problem.

Additional Notation: Let e_i be the i^{th} real standard basis vector, of dimension appropriate to the context in which it is used. For a matrix $X \in \mathbf{R}^{m\times n}$, let $X^\top$ be its transpose. Let $\mathbf{R}_+$ denote the set of nonnegative scalars. A set $\mathcal{C} \subseteq \mathbf{R}^n$ is a convex cone if it contains zero, and for any positive scalars a, b and any two vectors $x, y \in \mathcal{C}$, the vector $ax + by \in \mathcal{C}$. The dual cone of a given cone $\mathcal{C} \subseteq \mathbf{R}^n$ is given by $\mathcal{C}^* = \{y \in \mathbf{R}^n \mid x^\top y \geq 0,\ \forall\ x \in \mathcal{C}\}$. Lastly, given any set $\mathcal{S}$ and any matrix X, we denote by $\Pi_{\mathcal{S}}(X) \coloneqq \text{argmin}_{Y\in\mathcal{S}} \|X - Y\|_F$ the projection of X onto $\mathcal{S}$. Here, $\|\cdot\|_F$ denotes the Frobenius norm.

6.2 Preliminaries

6.2.1 Uncertainty Model

The uncertainty set $\Xi \subset \mathbf{R}^k$ is assumed to be a convex compact set given by

$$\Xi \ \coloneqq\ \{\xi \in \mathbf{R}^k \mid \xi_1 = 1,\ Q\xi \in \mathcal{C}\}, \tag{6.1}$$

where $Q \in \mathbf{R}^{\ell\times k}$, and $\mathcal{C} \subseteq \mathbf{R}^\ell$ is a proper cone. The requirement that $\xi_1 = 1$ is for notational convenience, as it enables the representation of affine functions of $(\xi_2, \ldots, \xi_k)^\top$ as linear functions of ξ. We assume that the linear hull of Ξ spans $\mathbf{R}^k$. Such assumption is without loss of generality since the dimension of ξ can be reduced, if necessary, to obtain an equivalent uncertainty set, which satisfies said assumption. We use this assumption in the proof of Lemma 6.2.1, which shows that a semi-infinite linear constraint can be equivalently represented by a finite number of conic constraints.

Before we state Lemma 6.2.1, it will be convenient to define the cone generated by the uncertainty set Ξ. It is given by

$$\text{cone}(\Xi) := \{\xi \in \mathbf{R}^k \mid \xi_1 \geq 0,\ Q\xi \in \mathcal{C}\}$$

The dual cone of $\text{cone}(\Xi)$ is given by

$$\text{cone}(\Xi)^* := \{\mu e_1 + Q^\top \lambda \mid \mu \in \mathbf{R}_+,\ \lambda \in \mathcal{C}^*\}. \tag{6.2}$$

Lemma 6.2.1. *Let $z \in \mathbf{R}^k$. Then, the following two statements are equivalent*

(i) $z^\top \xi \geq 0,\ \ \forall\, \xi \in \Xi,$

(ii) $z \in \text{cone}(\Xi)^*.$

The proof of Lemma 6.2.1 is omitted, as it relies on a simple duality argument that is central to the robust optimization paradigm (cf. Theorem 1.3.4 in [6]).

6.2.2 Exact and Approximate Solutions to RSDPs

In this section, we review some results from the literature that provide conditions on the structure of the uncertainty set Ξ, which enable the tractable reformulation of the RSDP.

First, for so-called *scenario-generated* uncertainty sets described by

$$\Xi = \text{conv}\,\{\xi^1, \ldots, \xi^N\}, \tag{6.3}$$

where conv S denotes the convex hull of S, the RSDP admits a reformulation as a semidefinite program. We have the following Theorem from [6, Chap. 8.1].

Theorem 6.2.2. *Consider a scenario-generated uncertainty Ξ of the form* (6.3)*. Then, the semidefinite program,*

$$\begin{aligned} \underset{x \in \mathbf{R}^m}{\text{minimize}} \quad & c^\top x \\ \text{subject to} \quad & x \in \mathbf{R}^m \\ & \sum_{i=1}^{k} \xi_i^j \mathcal{A}_i(x) \in \mathbf{S}_+^n, \quad j = 1, \dots, N, \end{aligned} \tag{6.4}$$

is equivalent to RSDP.

Second, given a $\rho \in \mathbf{R}$, *norm-bounded* uncertainty sets described by

$$\Xi = \{\xi \in \mathbf{R}^k \mid \xi = (1,\delta^1;\dots;\delta^N),\ \delta^j \in \mathbf{R}^{n_j}, \|\delta^j\|_2 \leq \rho,\ \forall\, j = 1,\dots,N\} \tag{6.5}$$

are known to yield tractable reformulations or (conservative) inner approximations for the RSDP. In particular, the authors in [6] show that RSDP admit a computationally tractable reformulation as a semidefinite program, if the uncertainty set is *unstructured* (i.e., $N = 1$). Under the more general setting of *structured* uncertainty (i.e., $N > 1$), a computationally tractable inner approximation to the RSDP can be derived. We summarize these results in the following Theorem, which is based on [88, Thm. 6.2.1] and [6, Thm 9.1.2 & Thm. 8.2.3].

Theorem 6.2.3. *Consider a norm-bounded uncertainty set Ξ of the form* (6.5)*. Let $\mu_0 := 1$.*

For each $j = 1, \ldots, N$, define $\mu_j := \sum_{s=1}^{j} n_s$ and

$$F_j(x, S_j, Q_j) := \begin{bmatrix} S_j & \rho\mathcal{A}_{\mu_{j-1}+1}(x) & \ldots & \rho\mathcal{A}_{\mu_j+1}(x) \\ \rho\mathcal{A}_{\mu_{j-1}+1}(x) & Q_j & & \\ \vdots & & \ddots & \vdots \\ \rho\mathcal{A}_{\mu_j+1}(x) & & & Q_j \end{bmatrix}.$$

(i) The semidefinite program

$$\begin{aligned} \text{minimize} \quad & c^\top x \\ \text{subject to} \quad & x \in \mathbf{R}^m, \ \{S_j\}_{j=1}^N, \ \{Q_j\}_{j=1}^N \in \mathbf{S}^n, \\ & F_j(x, S_j, Q_j) \succeq 0, \quad j = 1, \ldots, N, \\ & 2\mathcal{A}_1(x) - \sum_{j=1}^{N}(S_j + Q_j) \succeq 0, \end{aligned} \tag{6.6}$$

is an inner approximation of the RSDP, i.e., the projection of the feasible set of (6.6) on the space of x variables is contained in the feasible set of the RSDP.

(ii) If $N = 1$, then problem (6.6) is equivalent to the RSDP.

Lastly, for general semi-algebraic uncertainty sets Ξ, Scherer and Hol [72] develop a method based on sum of squares optimization to approximate the robust semidefinite program RSDP from within by a semidefinite program [1] We refer the reader to Theorem 1 of [72] for the details of their construction.

[1]In fact, their method method applies to the more general setting in which the robust constraint in the RSDP is polynomially parameterized in the uncertain parameter ξ.

6.3 Approximations of RSDPs

6.3.1 Outer Approximations

In this section, we propose a method for constructing tractable outer approximations to the robust semidefinite program RSDP. Consider the following *outer robust linear program* (O-RLP)

$$\begin{aligned} &\text{minimize} \quad c^\top x \\ &\text{subject to} \quad \sum_{i=1}^{k} \xi_i \mathcal{A}_i(x) \in \mathbf{P}, \quad \forall\, \xi \in \Xi, \end{aligned} \tag{O-RLP}$$

in the variables $x \in \mathbf{R}^m$, where $\mathbf{P}$ is an arbitrary polyhedral cone, that is assumed to contain the positive semidefinite cone. It is defined by

$$\mathbf{P} := \bigcap_{j=1}^{p} \{X \in \mathbf{S}^n \mid \mathrm{tr}(Z_j X) \geq 0\}, \tag{6.7}$$

where $Z_j \in \mathbf{S}^n_+$ for all $j = 1, \ldots, p$. Henceforth, we refer to $\mathbf{P}$ as the *outer polyhedral cone.* One example of an outer polyhedral cone is the cone of symmetric matrices with positive diagonal entries. It is obtained by taking $Z_i = e_i e_i^\top$, for all $i = 1, \ldots, n$. Since $\mathbf{P} \supseteq \mathbf{S}^n_+$, it follows that the optimal value of the O-RLP stands as a *lower bound* on the optimal value of the RSDP.

In what follows, we use Lemma 6.2.1 to obtain an equivalent finite-dimensional reformulation of the O-RLP. We have the following Proposition. Recall that v^{opt} denotes the optimal value of the RSDP.

Proposition 6.3.1. *The robust linear program O-RLP admits an equivalent reformulation as* *the following finite-dimensional conic linear program*

$$\begin{aligned}
\underset{x \in \mathbf{R}^m}{\text{minimize}} \quad & c^\top x && (6.8)\\
\text{subject to} \quad & x \in \mathbf{R}^m \\
& \sum_{i=1}^{k} \operatorname{tr}\big(Z_j \mathcal{A}_i(x)\big) e_i \in \operatorname{cone}(\Xi)^*, \quad j = 1, \ldots, p.
\end{aligned}$$

Let ℓ^{opt} denote the optimal value of the above program. It holds that $\ell^{\text{opt}} \leq v^{\text{opt}}$.

6.3.2 Inner Approximations

In this section, we propose one approach for constructing tractable inner approximations to the RSDP. In a similar vein to Section 6.3.1, we approximate the positive semidefinite cone from within by a *polyhedral cone*. Henceforth, we refer to this cone as an *inner polyhedral cone*.

Arbitrary inner polyhedral cones can be constructed by taking the conic hull of a finite number of positive semidefinite matrices. For simplicity, we choose the cone, which is dual to the outer polyhedral cone $\mathbf{P}$, defined in Section 6.3.1. It is given by

$$\mathbf{P}^* = \operatorname{cone}\{Z_1, \ldots, Z_p\} = \left\{ \sum_{j=1}^{p} y_j Z_j \;\middle|\; y \geq 0 \right\}. \tag{6.9}$$

Consider the following *inner robust linear program* (I-RLP)

$$\begin{aligned}
\underset{x \in \mathbf{R}^m}{\text{minimize}} \quad & c^\top x && \text{(I-RLP)}\\
\text{subject to} \quad & \sum_{i=1}^{k} \xi_i \mathcal{A}_i(x) \in \mathbf{P}^*, \quad \forall\, \xi \in \Xi,
\end{aligned}$$

Since I-RLP is an inner approximation to the RSDP, it follows that any feasible solution to the I-RLP will be feasible for the RSDP. In addition, the optimal value of the I-RLP serves as an *upper bound* on the optimal value of the RSDP.

We proceed with an an equivalent reformulation of the robust linear program I-RLP, which will prove useful in the sequel. Using the V(ertex)-representation of the polyhedral cone $\mathbf{P}^*$, it is straightforward to show that the robust linear program I-RLP can be exactly reformulated as the infinite-dimensional program

$$\begin{array}{ll} \underset{x \in \mathbf{R}^m,\ \{\mathsf{y}_j\}_{j=1}^p \in \mathbf{L}_{k,1}}{\text{minimize}} & c^\top x \\ \text{subject to} & \left.\begin{array}{l} \displaystyle\sum_{i=1}^k \xi_i \mathcal{A}_i(x) = \sum_{j=1}^p \mathsf{y}_j(\xi) Z_j, \\ \mathsf{y}_j(\xi) \geq 0, \quad j = 1, \ldots, p, \end{array}\right\} \forall\, \xi \in \Xi. \end{array} \tag{6.10}$$

Here, $\mathbf{L}_{k,1}$ denotes the infinite-dimensional space of all functions from $\mathbf{R}^k$ to $\mathbf{R}$. The above optimization problem is intractable in general as it involves infinite-dimensional decision variables and semi-infinite *equality* constraints. The difficulty derives from the V-representation of the polyhedral cone $\mathbf{P}^*$, which precludes a direct application of Lemma 6.2.1 to the semi-infinite constraint in (6.10). However, the Weyl-Minkowski theorem [13, Theorem 3.2] ensures that the polyhedral cone $\mathbf{P}^*$ has an equivalent H(yperplane)-representation of the following form

$$\mathbf{P}^* = \bigcap_{j=1}^{l} \{X \in \mathbf{S}^n \mid \mathrm{tr}(H_j X) \geq 0\},$$

for some finite natural number l. The H-representation of $\mathbf{P}^*$ can be used together with Lemma 6.2.1 to obtain an equivalent reformulation of the I-RLP as a finite-dimensional conic linear program. A practical drawback of this approach, however, is that it requires a preprocessing step to obtain a H-representation of the polyhedral cone $\mathbf{P}^*$

from its V-representation. Besides being computationally demanding to compute, an H-representation of the polyhedral cone may require a large number of hyperplanes for its specification. This, in turn, gives rise to optimization problems, which have a large number of constraints, thereby jeopardizing their tractability. This practical limitation raises the question as to whether it is possible to work directly with the vertex representation of the polyhedral cone $\mathbf{P}^*$ to obtain an inner approximation of the I-RLP that is less computationally demanding to solve. In what follows, we propose one such approach.

To obtain a tractable inner approximation of (6.10), we restrict the functional form of the functions $y_j,\ j = 1, \ldots, p$, to be linear in the uncertain parameter ξ. This approach is similar in spirit to the use of affine decision rules to approximate the infinite-dimensional decision space in stochastic programs [7]. More precisely, for each $j = 1, \ldots, p$, we require that

$$\mathsf{y}_j(\xi) \ = \ y_j^\top \xi,$$

for some vector $y_j \in \mathbf{R}^k$. This gives rise to the following *restricted inner robust linear program* (RI-RLP), which amounts to an inner approximation of the original I-RLP.

$$\begin{aligned}
&\underset{x\in\mathbf{R}^m,\ \{y_j\}_{j=1}^p\in\mathbf{R}^k}{\text{minimize}} \quad c^\top x && \text{(RI-RLP)}\\
&\text{subject to} \quad \left.\begin{aligned}&\sum_{i=1}^k \xi_i\Big(\mathcal{A}_i(x) - \sum_{j=1}^p y_j^\top e_i Z_j\Big) = 0,\\ &y_j^\top\xi \geq 0, \quad j = 1, \ldots, p,\end{aligned}\right\}\ \forall\, \xi \in \Xi.
\end{aligned}$$

We now develop an equivalent reformulation of the (restricted) robust linear program RI-RLP as a finite-dimensional conic linear program. First notice that, since the semi-infinite equality constraint in RI-RLP must hold for all $\xi \in \Xi$, the linear hull of Ξ must be contained in the nullspace of the linear map

$$\xi \mapsto \sum_{i=1}^{k} \xi_i \Big(\mathcal{A}_i(x) - \sum_{j=1}^{p} y_j^\top e_i Z_j \Big). \tag{6.11}$$

Moreover, since Ξ is assumed to span all of $\mathbf{R}^k$, the equality constraint holds if and only if the expression inside the parentheses in (6.11) is equal to zero for all $i = 1, \dots, k$. Finally, a direct application of Lemma 1 to the remaining semi-infinite inequality constraints in RI-RLP yields the finite-dimensional conic linear program in Proposition 6.3.2, which is an equivalent reformulation of the robust linear program (RI-RLP). Recall that v^{opt} denotes the optimal value of the RSDP.

Proposition 6.3.2. *The robust linear program RI-RLP admits an equivalent reformulation as a finite-dimensional conic linear program, given by*

$$\begin{array}{lll} \underset{x \in \mathbf{R}^m,\ \{y_j\}_{j=1}^p \in \mathbf{R}^k}{\text{minimize}} & c^\top x & \\ \text{subject to} & x \in \mathbf{R}^m,\ \{y_j\}_{j=1}^p \in \mathbf{R}^k & \\ & \mathcal{A}_i(x) = \sum_{j=1}^{p} y_j^\top e_i Z_j, & i = 1, \dots, k, \\ & y_j \in \text{cone}(\Xi)^*, & j = 1, \dots, p, \end{array} \tag{6.12}$$

Let u^{opt} be the optimal value and $(x^\star, \{y_j^\star\}_{j=1}^p)$ an optimal solution of the above program. Then,

(i) $u^{\text{opt}} \geq v^{\text{opt}}$ *and*

(ii) $x^\star$ *is a feasible solution for the RSDP.*

The key results from the previous two sections are summarized in the following Theorem.

Theorem 6.3.3. *The optimal value of the RSDP is bounded by*

$$\ell^{\text{opt}} \leq v^{\text{opt}} \leq u^{\text{opt}}$$

6.4 Recursive Approximations via Cutting Planes

In the previous section, we proposed a method for constructing outer and inner approximations to the RSDP. The effectiveness of these approximations, however, depends critically on the choice of the polyhedral cone $\mathbf{P}$, and for any given problem it is unclear what the best choice for $\mathbf{P}$ is. A naive approach might entail the construction of a hierarchy of inner and outer polyhedral cones via a uniform discretization of the boundary of the cross polytope [46]. For high levels in the hierarchy, however, this approach can create computational inefficiencies due to the large number of half-spaces (positive semidefinite matrices) defining the outer (inner) polyhedral cone. These, in turn, yield finite-dimensional programs (6.8) and (6.12) with a large number of variables and constraints.

In this paper, we build upon the previous approach by exploring polyhedral approximations of the positive semidefinite cone that are adaptively guided by the objective function. More precisely, starting with a coarse outer polyhedral cone (i.e., the cone of symmetric matrices with nonnegative diagonal entries), we prescribe a recursive method to refine this cone, which uses the solution at the current iteration step. Specifically, at each iteration step, we identify a number of matrices from a primal-dual optimal solution pair to the inner and outer robust linear programs, which are shown to lie on the boundary of the outer polyhedral cone. We project these matrices onto the positive semidefinite cone and refine the outer polyhedral cone by intersecting it with the half-spaces corresponding to the supporting hyperplanes at said projection points. The resulting feasible set of the outer (inner) approximation of the RSDP over the refined cone is shown to exclude the primal (dual) optimal solution at the previous iteration step. In Appendix B.8.1, we present a variant of this recursive method, which can be used to improve the performance of inner and outer polyhedral approximations to semidefinite programs.

In what follows, we assume, without loss of generality, that for each $i \in \{1, \ldots, k\}$, the affine function $\mathcal{A}_i : \mathbf{R}^m \to \mathbf{S}^n$ is parameterized by matrices $A_{ij} \in \mathbf{S}^n$, $j = 0, 1, \ldots, m$ as follows:

$$\mathcal{A}_i(x) := A_{i0} + \sum_{j=1}^{m} x_j A_{ij}.$$

6.4.1 Outer Approximations

In this section, we develop a recursive method, which yields a sequence of outer approximations to the RSDP. At each iteration step, our method uses an optimal solution of the O-RLP to refine the outer polyhedral cone $\mathbf{P}$ only in those regions that are important for optimization.

Given a feasible solution $x \in \mathbf{R}^m$ to the O-RLP, we define, as the initial step in the development of the recursive method, the following set of matrices

$$\mathcal{O}(x) := \left\{ \sum_{i=1}^{k} \xi_i \mathcal{A}_i(x) \;\middle|\; \xi \in \Xi \right\}. \tag{6.13}$$

Since x is a feasible solution to the O-RLP, it follows that $\mathcal{O}(x)$ must be contained in $\mathbf{P}$. In addition, the set $\mathcal{O}(x)$ will contain matrices that are not positive semidefinite, if x is not in the feasible set of the RSDP. In this case, there must exist a supporting hyperplane to the positive semidefinite cone such that $\mathcal{O}(x)$ is not contained in the corresponding half-space containing $\mathbf{S}^n_+$. We call such a hyperplane a *cutting plane* for $\mathcal{O}(x)$. More precisely, a cutting plane for $\mathcal{O}(x)$ is defined to be a hyperplane characterized by a matrix $Z \in \mathbf{S}^n_+$ and passing through the origin, such that

$$\mathcal{O}(x) \nsubseteq \{X \in \mathbf{S}^n \mid \mathrm{tr}(ZX) \geq 0\}. \tag{6.14}$$

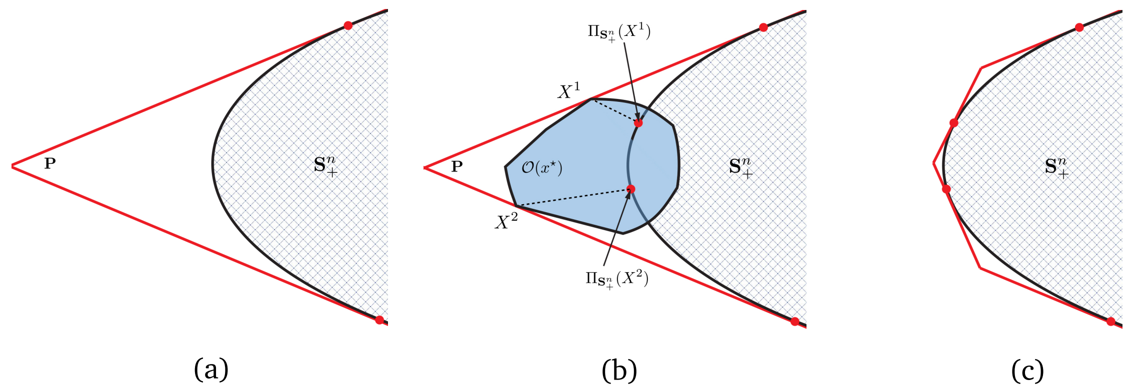


Figure 6.1: Figure 6.1a shows the positive semidefinite cone and the outer polyhedral cone $\mathbf{P}$, which is described by two half-spaces. In Figure 6.1b, we visualize the set $\mathcal{O}(x^\star)$ induced by an optimal solution $x^\star$ to the O-RLP. We identify two matrices, X^1 and X^2 in $\mathcal{O}(x^\star)$ that lie on the boundary of $\mathbf{P}$. The supporting hyperplanes to $\mathbf{S}^n_+$ at $\Pi_{\mathbf{S}^n_+}(X^1)$ and $\Pi_{\mathbf{S}^n_+}(X^2)$ cut off parts of $\mathcal{O}(x^\star)$ that do not contain positive semidefinite matrices. In Figure 6.1c, we visualize the refined outer polyhedral cone obtained by intersecting $\mathbf{P}$ with the half-spaces corresponding to the supporting hyperplanes. As seen in Proposition 6.4.1, the O-RLP over the refined cone does not include $x^\star$ in its feasible set.

An important consequence of (6.14) is that the O-RLP, which arises from approximating the positive semidefinite cone by $\mathbf{P} \cap \{X \in \mathbf{S}^n \mid \text{tr}(ZX) \geq 0\}$ will not contain the vector x in its feasible set. Recall the functional form of the constraints in problem (6.8). We have the following Proposition.

Proposition 6.4.1. *Let x be a feasible solution to the O-RLP. If*

$$\{X \in \mathbf{S}^n \mid \text{tr}(ZX) = 0\},$$

is a cutting plane for $\mathcal{O}(x)$, then

$$\sum_{i=1}^{k} \text{tr}\big(Z\mathcal{A}_i(x)\big)e_i \notin \text{cone}(\Xi)^*.$$

Proof. Since

$$\mathcal{O}(x^\star) \nsubseteq \{X \in \mathbf{S}^n \mid \text{tr}(ZX) \geq 0\},$$

it follows that there exists $\xi \in \Xi$ such that

$$0 > \sum_{i=1}^{k} \xi_i \text{tr}(Z\mathcal{A}_i(x^\star)) = \left(\sum_{i=1}^{k} \text{tr}(Z\mathcal{A}_i(x^\star))e_i\right)^\top \xi.$$

Since $\xi \in \Xi$, it follows that $\xi \in \text{cone}(\Xi)$. The desired exclusion follows from the definition of the dual cone. ■

Given an optimal solution $x^\star$ to the O-RLP, the above Proposition implies that a cutting plane for $\mathcal{O}(x^\star)$ yields an outer polyhedral cone such that the O-RLP over this cone does not include $x^\star$ in its feasible set. In what follows, we propose a technique for constructing cutting planes for $\mathcal{O}(x^\star)$. To describe our method, we need to consider the dual problem of the finite-dimensional conic linear program (6.8). It is given by

$$\begin{aligned}
\underset{\{z_j\}_{j=1}^p \in \mathbf{R}^k}{\text{maximize}} \quad & -\sum_{j=1}^{p}\sum_{t=1}^{k} \text{tr}(A_{i0}Z_j) z_j^\top e_t, && \\
\text{subject to} \quad & c_i = \sum_{j=1}^{p}\sum_{t=1}^{k} \text{tr}(A_{ti}Z_j) e_t^\top z_j, && i = 1, \ldots, m, \\
& z_j \in \text{cone}(\Xi), && j = 1, \ldots, p.
\end{aligned} \tag{6.15}$$

Using an optimal solution to (6.15), we identify a set of matrices in $\mathcal{O}(x^\star)$ that lie on the boundary of $\mathbf{P}$ (see Figure 6.1b). We have the following Lemma, whose proof follows from complementary slackness of conic duality and is omitted for brevity.

Lemma 6.4.2. *Let $(x^\star, \{z_j^\star\}_{j=1}^p)$ be a primal-dual optimal solution pair for programs* (6.8) *and* (6.15)*. For each index $j = 1, \ldots, p$, for which $(z_j)_1 > 0$, the following properties hold:*

(i) The vector $\xi^j = z_j^\star/(z_j^\star)_1 \in \Xi$.

(ii) The matrix $X^j \in \mathcal{O}(x^\star)$, defined by

$$X^j := \sum_{i=1}^{k} \xi_i^j \mathcal{A}_i(x^\star) \tag{6.16}$$

is on the hyperplane $\{X \in \mathbf{S}^n \mid \mathrm{tr}(Z_j X) = 0\}$.

Fix an index $j \in \{1, \ldots, p\}$. The following Proposition shows that if X^j is not positive semidefinite, then the supporting hyperplane to $\mathbf{S}^n_+$ at $\Pi_{\mathbf{S}^n_+}(X^j)$ is a cutting plane for $\mathcal{O}(x^\star)$.

Proposition 6.4.3. *Let* $(x^\star, \{z_j^\star\}_{j=1}^p)$ *be a primal-dual optimal solution pair to programs* (6.8) *and* (6.15)*. For each* $j = 1, \ldots, p$*, let* $X^j \in \mathcal{O}(x^\star)$ *be a matrix defined as in* (6.16)*. If* $X^j \notin \mathbf{S}^n_+$*, then*

$$\{X \in \mathbf{S}^n \mid \mathrm{tr}\big((\Pi_{\mathbf{S}^n_+}(X^j) - X^j)X\big) = 0\}.$$

is a cutting plane for $\mathcal{O}(x^\star)$.

Proof. Since $X^j \in \mathcal{O}(x^\star)$, it suffices to show that

$$X^j \notin \{X \in \mathbf{S}^n \mid \mathrm{tr}\big((\Pi_{\mathbf{S}^n_+}(X^j) - X^j)X\big) \geq 0\}, \tag{6.17}$$

as this implies that

$$\mathcal{O}(x^\star) \nsubseteq \{X \in \mathbf{S}^n \mid \mathrm{tr}\big((\Pi_{\mathbf{S}^n_+}(X^j) - X^j)X\big) \geq 0\}.$$

Let $X^j = \sum_{t=1}^n \lambda_t u_t u_t^\top$ be an eigenvalue decomposition of X^j. It follows that

$$\Pi_{\mathbf{S}^n_+}(X^j) - X^j = -\sum_{t=1}^n \min\{0, \lambda_t\} u_t u_t^\top.$$

Since $X^j \notin \mathbf{S}^n_+$, it follows that

$$\begin{aligned}
0 &> \sum_{t=1}^n - \min\{\lambda_t, 0\}\lambda_t \\
&= \sum_{t=1}^n - \min\{\lambda_t, 0\} u_t^\top X^j u_t \\
&= \sum_{t=1}^n - \min\{\lambda_t, 0\} u_t^\top \left(\sum_{i=1}^k \xi_i^j \mathcal{A}_i(x)\right) u_t \\
&= \sum_{i=1}^k \xi_i^j \mathrm{tr}\left((\Pi_{\mathbf{S}^n_+}(X^i) - X^i)\mathcal{A}_i(x)\right),
\end{aligned}$$

which establishes the desired exclusion (6.17). ■

Some remarks regarding Proposition 6.4.3 are in order. First, the cutting planes defined in Proposition 6.4.3 *cut off* a parts of $\mathcal{O}(x^\star)$ that do not contain positive semidefinite matrices (see Figure 6.1b). Second, the specification of the cutting planes in Proposition 6.4.3 follows from the characterization of nearest points to nonempty closed convex sets (see, for example, Excercise 8(c) in [12]). Third, as shown in Proposition 6.4.1, *for any* index $j \in \{1, \ldots, p\}$ for which $X^j \notin \mathbf{S}^n_+$, the outer polyhedral cone

$$\mathbf{P} \cap \{X \in \mathbf{S}^n \mid \mathrm{tr}\big((\Pi_{\mathbf{S}^n_+}(X^j) - X^j)X\big) \geq 0\Big\}$$

yields an outer approximation to the robust semidefinite program, which does not contain $x^\star$ in its feasible set. Clearly, there is a computational tradeoff between the number of cutting planes one chooses for refining $\mathbf{P}$ (the larger the number of cutting planes, the more the constraints of the resulting O-RLP) and the marginal improvement to the lower bound obtained from the resulting outer approximation. In Figure 6.1c, we visualize the outer polyhedral cone, which is obtained by intersecting $\mathbf{P}$ with *all* the corresponding half-spaces. Lastly, if $x^\star$ is uniquely optimal, then the optimal value of the O-RLP over the refined outer polyhedral cone is guaranteed to be greater than the optimal value of the corresponding approximation over $\mathbf{P}$.

In Table 6.1, we outline the steps of the proposed algorithm. Starting with a coarse outer polyhedral cone $\mathbf{P}$, the algorithm computes an optimal solution $x^\star$ to the O-RLP by solving the finite-dimensional conic linear program (6.8). A number of cutting planes for $\mathcal{O}(x^\star)$ is then computed according to Proposition 6.4.3. The intersection of $\mathbf{P}$ with the corresponding half-spaces yields a refined outer polyhedral cone. The above procedure is repeated using the refined cones until all the matrices in Proposition 6.4.3 are positive semidefinite

Algorithm : Recursive Outer Approximations of RSDP

Given an outer polyhedral cone $\mathbf{P}$ and a maximum number $\bar{t}$ of iterations
Initialize $t = 1$
Repeat

1. *Let* $p =$ number of half-spaces defining $\mathbf{P}$

2. *Compute.*

   - A primal-dual optimal solution pair to (6.8) and (6.15)
   - Matrices $X^1, \ldots, X^p$ defined according to (6.16)

3. *Update.*

   - $\mathbf{P} = \mathbf{P} \bigcap_{j=1}^{p} \{X \in \mathbf{S}^n \mid \mathrm{tr}\big((\Pi_{\mathbf{S}^n_+}(X^j) - X^j)X\big) \geq 0\}$
   - $t = t + 1$

Until $X^j \in \mathbf{S}^n_+$, for all $j = 1, \ldots, p$ or $t = \bar{t}$
Output The optimal value ℓ^{opt} of problem (6.8)

Table 6.1: Recursive cutting plane algorithm, which yields a sequence of outer approximations to the RSDP, whose costs is nondecreasing.

or until a maximum number of iterations is reached.

6.4.2 Inner Approximations

In this section, we develop a recursive method, which yields a sequence of inner approximations to the RSDP. Contrary to Section 6.4.1, the method we develop in this section is applied to the dual problem of the finite-dimensional conic program (6.12) of the RI-RLP.

This is given by

$$\begin{aligned} \text{maximize} \quad & -\sum_{t=1}^{k} \operatorname{tr}(S_t A_{t0}) \\ \text{subject to} \quad & c_i = \sum_{t=1}^{k} \operatorname{tr}(S_t A_{ti}), \qquad & i = 1, \ldots, m \\ & \sum_{t=1}^{k} \operatorname{tr}(Z_j S_t) e_t \in \operatorname{cone}(\Xi), \qquad & j = 1, \ldots, p, \end{aligned} \tag{6.18}$$

in the variables $\{S_t\}_{t=1}^k \in \mathbf{S}^n$. Our method uses an optimal solution of (6.18) to refine the outer polyhedral cone $\mathbf{P}$ only in those regions that are important for optimization. As seen from the last set of constraints in (6.18), such refinements of $\mathbf{P}$ shrink the feasible set of (6.18), thereby enlarging the feasible set of (6.12). This, in turn, yields a sequence of inner approximation whose costs in nonincreasing. We make the following assumption, which ensures that the optimal value of (6.18) is an upper bound to the optimal value of the RSDP.

Assumption 6.4.4. *Strong duality holds between the primal program* (6.12) *and its dual* (6.18).

Given a feasible solution $\{S_i\}_{i=1}^k \in \mathbf{S}^n$ to (6.18) we define, as the initial step in the construction of the recursive method, the following set of matrices

$$\mathcal{I}(S_1, \ldots, S_k) := \left\{ \sum_{i=1}^{k} y_i S_i \;\middle|\; y \in \operatorname{cone}(\Xi)^* \right\}. \tag{6.19}$$

In the following Lemma, we collect some important properties of $\mathcal{I}(S_1, \ldots, S_k)$.

Lemma 6.4.5. *The following properties hold:*

(i) Let $\{S_i\}_{i=1}^k \in \mathbf{S}^n$ be a feasible solution to program (6.18)*. Then,*

$$\mathcal{I}(S_1, \ldots, S_k) \subseteq \mathbf{P}.$$

In addition, $\mathcal{I}(S_1,\dots,S_k) \subseteq \mathbf{S}^n_+$, if and only if

$$\sum_{t=1}^{k} \operatorname{tr}(ZS_t)e_t \in \operatorname{cone}(\Xi), \quad \forall\, Z \in \mathbf{S}^n_+. \tag{6.20}$$

(ii) Let $(\{y_j^\star\}_{j=1}^p, \{S_i^\star\}_{i=1}^k,)$ be a primal-dual optimal solution pair to programs (6.12) and (6.18). For each $j = 1,\dots.p$, the matrix $S^j \in \mathcal{I}(S_1^\star,\dots,S_k^\star)$, defined by

$$S^j := \sum_{i=1}^{k} (y_j^\star)_i S_i^\star, \tag{6.21}$$

is on the hyperplane $\{X \in \mathbf{S}^n \mid \operatorname{tr}(Z_jX) = 0\}$.

Remark 14. *If the cone $\mathcal{C}$ characterizing the uncertainty set Ξ is equal to the positive orthant, then (6.20) can be equivalently represented by a set of linear matrix inequalities. In particular, $\mathcal{I}(S_1,\dots,S_k) \subseteq \mathbf{S}^n_+$ if and only if*

(i) $S_1 \in \mathbf{S}^n_+$, and

(ii) $\displaystyle\sum_{t=1}^{k} e_i^\top Q e_t S_t \in \mathbf{S}^n_+$, for all $i = 1,\dots,k$.

The last constraint in (6.18) and the necessary and sufficient condition (6.20) point in the direction of a recursive method, for refining the outer polyhedral. Namely, we would like to find a supporting hyperplane to the positive semidefinite cone such that $\mathcal{I}(S_1,\dots,S_k)$ is not contained in the corresponding half-space containing $\mathbf{S}^n_+$. We call such a hyperplane a *cutting plane* for $\mathcal{I}(S_1,\dots,S_k)$. More precisely, a cutting plane for $\mathcal{I}(S_1,\dots,S_k)$ is defined to be a hyperplane characterized by a matrix $Z \in \mathbf{S}^n_+$ and passing through the origin, such that

$$\mathcal{I}(S_1,\dots,S_k) \nsubseteq \{X \in \mathbf{S}^n \mid \operatorname{tr}(ZX) \geq 0\}. \tag{6.22}$$

An important consequence of (6.22) is that the dual program (6.18), which is obtained from approximating the positive semdefinite cone by $\mathbf{P} \cap \{X \in \mathbf{S}^n \mid \operatorname{tr}(ZX) \geq 0\}$ will not

Algorithm: Recursive Inner Approximations of RSDP

Given an outer polyhedral cone $\mathbf{P}$ and a maximum number $\bar{t}$ of iterations
Initialize $t = 1$
Repeat

1. *Let* $p =$ number of half-spaces defining $\mathbf{P}$
2. *Compute.*
   - A primal-dual optimal solution pair $(x^\star, \{y_j^\star\}_{j=1}^p, \{S_i^\star\}_{i=1}^k)$ to programs (6.12) and (6.18)
   - Matrices $\{S^j\}_{j=1}^p$ defined according to (6.16)
3. *Update.*
   - $\mathbf{P} = \mathbf{P} \bigcap_{j=1}^p \{X \in \mathbf{S}^n \mid \text{tr}\big((\Pi_{\mathbf{S}_+^n}(S^j) - S^j)X\big) \geq 0\}$
   - $t = t + 1$

Until $S^j \in \mathbf{S}_+^n$, for all $j = 1, \ldots, p$ or $t = \bar{t}$
Output The optimal value u^{opt} and an optimal solution $x^\star \in \mathbf{R}^m$ to problem (6.12).

Table 6.2: Recursive cutting plane algorithm, which yields a sequence of outer approximations to the RSDP, whose costs is nondecreasing.

contain $\{S_t\}_{t=1}^k$ in its feasible set. We have the following Proposition whose proof relies on similar arguments as the proof of Proposition 6.4.1 and is therefore omitted.

Proposition 6.4.6. *Let $\{S_i\}_{i=1}^k$ be a feasible solution to (6.18). If*

$$\{X \in \mathbf{S}^n \mid \text{tr}(ZX) = 0\}$$

is a cutting plane for $\mathcal{I}(S_1, \ldots, S_k)$, then

$$\sum_{t=1}^k \text{tr}(ZS_t)e_t \notin \text{cone}(\Xi)$$

Given a primal-dual optimal solution pair $(\{y_j^\star\}_{j=1}^p, \{S_i^\star\}_{i=1}^k)$ to programs (6.12) and (6.18), we propose a method for constructing cutting planes for $\mathcal{I}(S_1^\star, \ldots, S_k^\star)$. Lemma 6.4.5((ii)) identifies a set of matrices, $\{S^j\}_{j=1}^p$ in $\mathcal{I}(S_1^\star, \ldots, S_k^\star)$, which lie on the boundary of the outer polyhedral cone $\mathbf{P}$. The following Proposition shows that if S^j is not positive semidefinite, then the supporting hyperplane to $\mathbf{S}_+^n$ at $\Pi_{\mathbf{S}_+^n}(S^j)$ is a cutting plane for $\mathcal{I}(S_1^\star, \ldots, S_k^\star)$.

Proposition 6.4.7. *Let $(\{y_j^\star\}_{j=1}^p, \{S_t^\star\}_{t=1}^k)$ be a primal-dual optimal solution pair to (6.12) and (6.18). For each $j = 1, \ldots, p$, let $S^j \in \mathcal{I}(S_1^\star, \ldots, S_k^\star)$ be a matrix defined as in (6.21). If $S^j \notin \mathbf{S}_+^n$, then*

$$\{X \in \mathbf{S}^n \mid \mathrm{tr}\big((\Pi_{\mathbf{S}_+^n}(S^j) - S^j)X\big) = 0\}.$$

is a cutting plane for $\mathcal{I}(S_1^\star, \ldots, S_k^\star)$.

Proof. Since $S^j \in \mathcal{I}(S_1^\star, \ldots, S_k^\star)$, it suffices to show that

$$S^j \notin \{X \in \mathbf{S}^n \mid \mathrm{tr}\big((\Pi_{\mathbf{S}_+^n}(S^j) - S^j)X\big) \geq 0\}, \tag{6.23}$$

as this implies that

$$\mathcal{I}(S_1^\star, \ldots, S_k^\star) \not\subseteq \{X \in \mathbf{S}^n \mid \mathrm{tr}\big((\Pi_{\mathbf{S}_+^n}(S^j) - S^j)X\big) \geq 0\}.$$

Let $S^j = \sum_{i=1}^n \lambda_i u_i u_i^\top$ be an eigenvalue decomposition of S^j. It follows that

$$\Pi_{\mathbf{S}_+^n}(S^j) - S^j = -\sum_{i=1}^n \min\{0, \lambda_i\} u_i u_i^\top.$$

Since $S^j \notin \mathbf{S}^n_+$, it follows that

$$\begin{aligned}
0 &> \sum_{i=1}^{n} -\min\{\lambda_i, 0\}\lambda_i \\
&= \sum_{i=1}^{n} -\min\{\lambda_i, 0\}u_i^\top X^j u_i \\
&= \sum_{i=1}^{n} -\min\{\lambda_i, 0\}u_i^\top S^j u_i \\
&= \sum_{i=1}^{n} -\min\{\lambda_i, 0\}u_i^\top \left(\sum_{t=1}^{k}(y_j)_t^\star S_t^\star\right) u_i \\
&= \operatorname{tr}\left((\Pi_{\mathbf{S}^n_+}(S^j) - S^j)\sum_{t=1}^{k}(y_j)_t^\star S_t^\star\right),
\end{aligned}$$

which establishes the desired exclusion. ■

In Table 6.2 , we outline the steps of the proposed algorithm, which is identical to the algorithm developed in Section 6.4.1.

6.5 An Application to Robust Linear Estimation

6.5.1 The Robust Linear Estimation Problem

We consider the robust linear estimation (RLE) problem studied in [6, 20]. In the RLE problem, a signal $x \in \mathbf{R}^n$ is observed through a linear measurement process

$$y = Ux + \varepsilon,$$

where $\varepsilon \in \mathbf{R}^m$ is a zero-mean random vector with known covariance matrix $\Sigma \in \mathbf{S}^m$ and $U \in \mathbf{R}^{m\times n}$ is an unknown matrix, which is known only up to a membership in a set

$$\mathcal{U} := \left\{ \sum_{i=1}^{k} \xi_i U_i \,\middle|\, \xi \in \Xi \right\},$$

where Ξ is a known uncertainty set of the form considered in (6.1). The signal x is also known to belong to a set

$$\mathcal{X} = \{x \in \mathbf{R}^n \mid x^\top R x \leq 1\},$$

where $R \in \mathbf{S}^n$ is a symmetric positive definite matrix. The objective in the RLE problem is to find a linear estimator

$$\hat{x} = Gy,$$

of x, defined by a matrix $G \in \mathbf{R}^{n\times m}$, which minimizes the worst-case root mean square error. More precisely, we aim to find a solution to the following optimization problem

$$\underset{G\in R^{n\times m}}{\text{minimize}} \quad \sup_{U\in\mathcal{U}, x\in\mathcal{X}} \sqrt{\mathbb{E}\left[\|\hat{x}-x\|_2^2\right]}, \tag{6.24}$$

where the expectation is over ε. As one can verify, the RLE problem (6.24) admits an equivalent reformulation as a robust semidefinite program of the following form:

$$\begin{array}{ll} \underset{t,\tau,\delta,G}{\text{minimize}} & t \\ \text{subject to} & \sqrt{\tau^2+\delta^2} \leq t \\ & \|G\Sigma^{1/2}\|_F \leq \delta \\ & \begin{bmatrix} \tau I & \mathcal{B}(G,\xi)^\top \\ \mathcal{B}(G,\xi) & \tau I \end{bmatrix} \in \mathbf{S}_+^{2n}, \ \forall\, \xi \in \Xi, \end{array} \tag{6.25}$$

in the variables $t, \tau, \delta \in \mathbf{R}$ and $G \in \mathbf{R}^{m\times n}$. Here,

$$\mathcal{B}(G,\xi) := -R^{-1/2} + \sum_{i=1}^{k} \xi_i R^{-1/2} G U_i.$$

The variables t, τ, δ arise because of the epigraphical formulation of the problem.

6.5.2 Numerical Studies

In what follows, we consider inner and outer polyhedral approximations to the robust semidefinite program RSDP and apply the cutting plane scheme developed in Section 6.4. The cutting plane method for the respective approximation terminates when all matrices X^j and S^j in Propositions 6.4.3 and 6.4.7 are positive semidefinite. We let $k = 6$, $n = 4$, and $m = 7$, and take the outer polyhedral cone $\mathbf{P}$ to be equal to

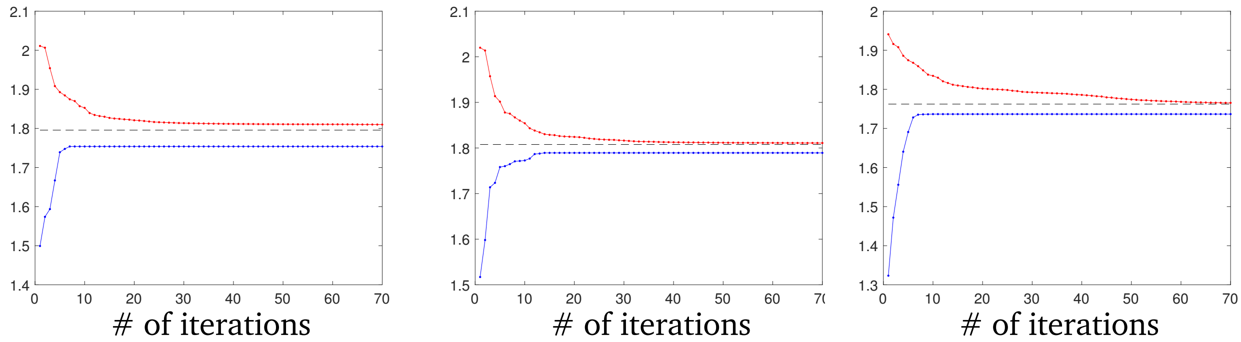
$$\mathbf{P} = \{X \in \mathbf{S}^8 \mid X_{jj} + X_{ii} + 2X_{ij} \geq 0,\ 1 \leq i, j \leq n\}.$$

The dual cone of $\mathbf{P}$ is the cone of diagonally dominant matrices with nonnegative diagonal entries. We generate matrices $U_1, \ldots, U_6 \in \mathbf{R}^{7\times 4}$ whose entries are sampled from the standard normal distribution and a random covariance matrix $\Sigma \in \mathbf{S}^7_+$, which is generated according to $\Sigma = PP^\top$, where $P \in \mathbf{R}^m$ is a random matrix whose entries are also drawn from the standard normal distribution. We consider three choices for the uncertainty set Ξ. At each iteration step, we apply the cutting plane method developed in Section 6.4, which ensures that the optimal solutions of the outer and inner approximations at the previous iteration step are excluded from the new feasible sets. We compare the performance of our approach with existing results from the literature [88, 6, 72], which are described in Section 6.2.2. All numerical analyses were carried out using SDPT3 [83].

(i) *Unstructured Normed-Bounded Uncertainty*. We consider first the case of unstructured norm-bounded uncertainty (cf. Section 6.2.1). More precisely, we let

$$\Xi = \{\xi \in \mathbf{R}^6 \mid \|\xi\|_2 \leq 2,\ \xi_1 = 1\}. \tag{6.26}$$

As shown in Theorem 6.2.3, this choice of Ξ yields a tractable reformulation of the RSDP as a semidefinite program. The optimal value of said semidefinite program is equal to 1.804. Figure 6.2(a), depicts the optimal values of the finite-dimensional



(a) (b) (c)

Figure 6.2: Optimal values of O-RLP and I-RLP at each iteration step of the algorithms in Tables 6.1 and 6.2, respectively. Figure 6.2(a) depicts the case of unstructured normed-bounded uncertainty (6.26), Figure 6.2(b) the case of structured normed-bounded uncertainty (6.27), and Figure 6.2(c) the case of polytopic uncertainty (6.28). All figures illustrate that the feasible solution to the RSDP obtained from the optimal solution to I-RLP at the termination of the inner approximation algorithm in Table 6.2 is nearly optimal.

outer and inner approximations O-RLP and RI-RLP, respectively at each iteration step. Each finite-dimensional program is a second-order cone program. The dashed line depicts the cost achieved by the sum of squares inner approximation method of [72]. The optimal value of the outer approximations converge to 1.754, achieving a 2.75% optimality gap. The optimal value of the inner approximation converge to 1.809, achieving a 0.3421% optimality gap. Notice that the sequence of optimal values associated with the outer and inner approximations nearly converge within the first few iterations.

(ii) *Structured Normed-Bounded Uncertainty*

Next, we consider the case of structured norm-bounded uncertainty (cf. Section 6.2.1). More precisely, we let

$$\Xi = \{\xi \in \mathbf{R}^6 \mid \|(\xi_2, \xi_3)\|_2 \leq 1,\ \|(\xi_4, \xi_5, \xi_6)\|_2 \leq 1,\ \xi_1 = 1\}. \tag{6.27}$$

Theorem 6.2.3 provides a conservative semidefinite program to calculate a feasible solution to the robust linear estimation problem given uncertainty sets with structured norm-bounded uncertainty sets. However, the resulting semidefinite program turns out to be infeasible for the specific problem being studied, and therefore it provides no useful information. Alternatively, the approximation method of [72] yields a feasible solution having cost equal to 1.808. This is depicted by the dashed line in Figure 6.2(b). The lower bound of 1.79 implied by the outer approximation algorithm we propose implies that this feasible solution is within a few percent of optimal. The inner polyhedral approximation scheme, we propose converges to 1.811 and is also with a few percent of optimal.

(iii) *Polytopic Uncertainty* Finally, we consider the case of polytopic uncertainty. More precisely, we let

$$\Xi \;=\; \{\xi \in \mathbf{R}^6 \mid \|\xi\|_\infty \leq 1,\; L\xi \geq 0,\; \xi_1 = 1\}. \tag{6.28}$$

Here, $L \in \mathbf{R}^{6\times 6}$ is a random matrix whose entries are sampled from the standard normal distribution and is given by

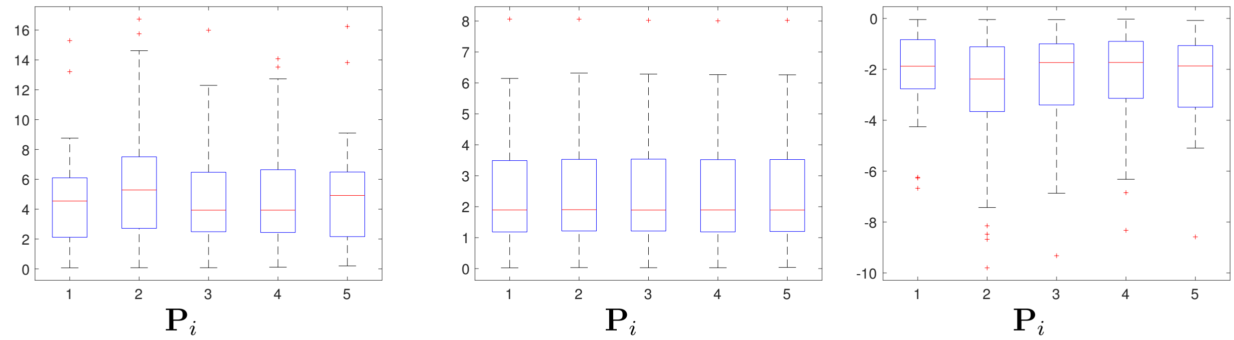
$$L = \begin{bmatrix} 0.9737 & -1.4916 & 0.7219 & -1.3222 & -1.1312 & -1.1312 \\ 0.2638 & 1.8082 & 0.0366 & -0.0893 & 1.2418 & -0.4643 \\ 0.0007 & -1.2102 & -0.1954 & 0.0412 & -0.1647 & -0.3545 \\ 0.2645 & 0.7456 & -0.4253 & 0.7794 & -1.3772 & 1.1866 \\ 0.6323 & 2.0034 & -1.0453 & -0.4640 & 0.9598 & -0.2211 \\ 0.3616 & -0.4560 & -1.4048 & -0.1552 & -0.1891 & 0.4325 \end{bmatrix}.$$

In Figure 6.2(c), we plot the optimal values of O-RLP and RI-RLP at each step of the corresponding algorithms presented in Tables 6.1 and 6.2. The finite-dimensional programs O-RLP and the RI-RLP are SOCPs. The optimal values of the outer approximation converge to 1.737, whereas the optimal values of the restricted inner approximation converge to 1.765. Therefore, the feasible solution computed achieves a cost, which is within 8.09% of the optimal value of the RSDP. The sum of squares inner approximation method of [72] returns a feasible solution which yields a cost of 1.762, achieving a slight improvement over our approach.

6.5.3 Sensitivity Analysis

In this section, we examine the performance of our method on several problem data and for several choices of the outer polyhedral cone $\mathbf{P}$ that is used to initialize the inner and outer approximation algorithms. In particular, we consider thirty different sets of data matrices $\{U_1, \ldots, U_6\}$ whose entries are drawn from the standard normal distribution. With each set of problem data, we associate a covariance matrix $\Sigma \in \mathbf{S}^m$, which is determined according to $\Sigma = PP^\top$. The entries of $P \in \mathbf{R}^{m \times m}$ are drawn from the standard normal distribution. A natural question that arises is the following: is the choice of the cone $\mathbf{P}$ critical for the values of the upper and lower bounds returned by the recursive algorithms? We present some experimental results addressing this question.

We consider five outer polyhedral cones $\mathbf{P}_i \supseteq \mathbf{S}^8_+$, $i = 1, \ldots, 5$, each of which is determined by sixty four rank-one matrices $v_j v_j^\top$, $j = 1, \ldots, 64$. The entries of the vectors $v_j \in \mathbf{R}^n,\ j = 1, \ldots, 64$ are uniformly distributed in the interval $[-1, 1]$. We disregard any polyhedral cones $\mathbf{P}_i$ that yield empty feasiblse regions for the inner approximation (6.12) of any of the thirty problem instances. For the purpose of this study, we focus on the structured-norm bounded uncertainty set in (6.27) and apply the outer and inner approximation algorithms in Tables 6.1 and 6.2, for each outer polyhedral cone $\mathbf{P}_i$ and each set of problem data.



(a) (b) (c)

Figure 6.3: Fig. 6.3(a) depicts the confidence intervals of the percent gap between the optimal value returned by the recursive inner and outer approximation algorithms we propose. Fig. 6.3(b) (Fig. 6.3(c)) depicts the confidence intervals of the percent gap between the the optimal value of the sum of squares inner approximation method in [72] and the inner (outer) approximation method developed in this chapter. The central line depicts the empirical median, the box shows the interquartile range, the whiskers extend to 1.5 times the interquartile range, and the crosses denote the outliers.

In Fig. 6.3(a), we plot the confidence intervals for the percent gap between the optimal value returned by the inner and outer recursive alogorithms across the thirty problem instances for each cone $\mathbf{P}_i$. The central line depicts the empirical median, the box shows the interquartile range, the whiskers extend to 1.5 times the interquartile range, and the crosses denote the outliers. Similarly, in Fig. 6.3(b) (Fig. 6.3(c)), we visualize the percent gap between the recursive inner (outer) approximation scheme we propose and the optimal value returned by the sum of squares inner approximation method in [72]. We observe that the optimal values of the inner and outer approximation methods are somewhat sensitive to the choice of the polyhedral cone $\mathbf{P}_i$. Among all polyhedral cones considered, $\mathbf{P}_2$ yields the largest percent gap between the optimal values of the recursive inner and outer approximation methods. Figures 6.3(b)-(c) also suggest that between the inner and outer approximation algorithms, $\mathbf{P}_2$ performs the worst on the latter. We also observe that for every problem instance, the inner approximation approach in [72] outperforms the recursive inner approximation algorithm we propose, but only by a small percentage.

We close this section by noting that we consider the problem of choosing an appropriate outer polyhedral cone $\mathbf{P}$ for initializing the recursive approximation algorithms to be a promising area for further theoretical investigation.

6.6 Conclusions

In this chapter, we investigated the problem of approximating solutions to intractable robust semidefinite programs. The method we propose relies on approximating the positive semidefinite cone from within and without by appropriate polyhedral cones. We also propose a recursive method, which refines said polyhedral cones only in those regions that are important for optimization. The practical value of the results derives from the fact that one can obtain a feasible solution to the RSDP by solving a finite-dimensional conic linear program; and can bound the suboptimality incurred by this feasible solution by solving another finite-dimensional conic linear program. An a posteriori certificate of near optimality of the feasible solution is obtained if the gap between the optimal values of the outer and inner approximations is small.

APPENDIX A

AC POWER FLOW MODEL

The problem of adjusting the operating conditions of generators to meet the ever-changing load demand is at the heart of power system operations. Although the demand of any particular unit can vary over time, the aggregate demand changes rather slowly. Therefore, within any small time period, the power system can be regarded as being in *steady-state*. The power flow equations, which are derived according to Kirchhoff's current and voltage laws, model the steady-state relationship between the complex bus voltages and power injections in an electric power network. In this section, we derive the steady state AC-Power Flow equations.

A.1 Steady State AC Power Flow

We consider an electric power network whose topology is described by an undirected graph $\mathcal{G} := (\mathcal{V}, \mathcal{E})$ without self-loops, where the vertex set $\mathcal{V} := \{1, \ldots, n\}$ represents the collection of network buses and the edge set $\mathcal{E}$ represents the collection of transmission lines connecting buses. We assume that $(i, j) \in \mathcal{E}$ if and only if $(j, i) \in \mathcal{E}$. Let $v_i \in \mathbf{C}$ be the complex power at bus i.

In power transmission networks, transmission lines are represented by the so-called nominal π-circuit model. The π-circuit model is illustrated in Fig. A.1(a) for a line connecting bus i and bus j. In this model, $y_{ij} \in \mathbf{C}$ denotes the series admittance of line (i, j) and $\widehat{y}_{ij} \in \mathbf{C}$ denotes the total shunt admittance. The total shunt admittance is usually modeled as two capacitors of equal value (the value of each capacitor is equal to half the total shunt

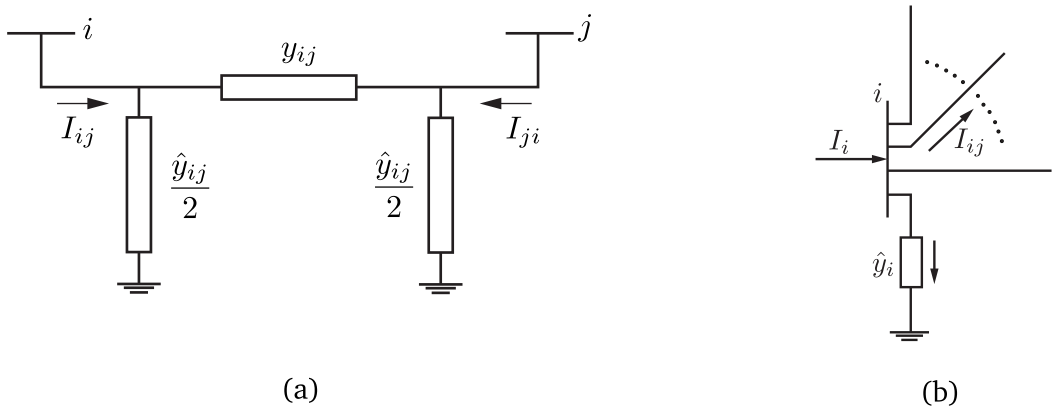


Figure A.1: (a) Transmission line π-circuit model for a line connecting buses i and j. (b) The net complex current injection at bus $i \in \mathcal{V}$, denoted by I_i, is equal to the sum of the current flows in the branches incident to bus i (i.e., I_{ij}, for $(i,j) \in \mathcal{E}$) plus the current flowing to ground through the total shunt admittance at bus i.

admittance), placed at the sending and receiving ends of the line. In addition, it is also assumed that $\mathrm{Re}(\hat{y}_{ij}) = 0$, that is the shunt conductance is equal to zero. The complex current from bus i to bus j is denoted by $I_{ij} \in \mathbf{C}$ and it is given by

$$I_{ij} = y_{ij}(v_i - v_j) + \frac{1}{2}\hat{y}_{ij}v_i.$$

where $v_i \in \mathbf{C}$ denote the complex voltage phasor at bus i. A similar expression can be derived for I_{ji}. Let

$$\hat{y}_i := \frac{1}{2} \sum_{j:(i,j)\in\mathcal{E}} \hat{y}_{ij},$$

denote the total shunt admittance at bus i. Kirchoff's current law states that the sum of all currents entering and leaving a node must be equal to zero, therefore, the net complex current injected at bus $i \in \mathcal{V}$ must equal

$$I_i \ = \hat{y}_i v_i + \sum_{j:(i,j)\in\mathcal{E}} y_{ij}(v_i - v_j).$$

(see Fig. A.1(b)). Letting $I \in \mathbf{C}^n$ be the vector of complex current injections and $v \in \mathbf{C}^n$ the vector of complex voltage phasors, the above set of equations can be expressed compactly as $I = Yv$, where $Y \in \mathbf{C}^{n\times n}$ is the *bus-admittance* matrix, given by

$$[Y]_{ij} := \begin{cases} \hat{y}_i + \sum\limits_{j:(i,j)\in\mathcal{E}} y_{ij}, & \text{if } i = j, \\ -y_{ij}, & \text{if } (i,j) \in \mathcal{E}, \\ 0, & \text{otherwise.} \end{cases} \tag{A.1}$$

The bus admittance matrix is a symmetric matrix. Moreover, its sparsity pattern captures the topology of the network, as y_{ij} is nonzero if and only if there a transmission line connecting buses i and j.

The complex power flow from bus i to bus j is defined as follows:

$$s_{ij} := v_i I_{ij}^* = |v_i|^2 \left(\frac{1}{2}\widehat{y}_{ij} + y_{ij}\right)^* + v_i v_j^* y_{ij}^*. \tag{A.2}$$

Moreover, the net complex power injection at bus $i \in \mathcal{V}$ is given by

$$s_i := v_i I_i^* = \sum_{j=1}^{n} v_i v_j^* [Y]_{ij}^*, \tag{A.3}$$

The set of equations (A.2) is known as the set of *AC power flow equations* and (A.3) as the set of *AC power balance equations*. The following Lemma provides a compact characterization for (A.2) and (A.5).

Lemma A.1.1 (AC Power Flow Equations)**.**

(i) The AC power balance equations are given by

$$s_i = v^* S_i v, \qquad i \in \mathcal{V}, \tag{A.4}$$

where $S_i := Y^* e_i e_i^\top \in \mathbf{C}^{n\times n}$ *and* $v \in \mathbf{C}^n$ *is the vector of complex bus voltages.*

(ii) The AC power flow equations are given by

$$s_{ij} = v^* S_{ij} v, \qquad (i,j) \in \mathcal{E}, \tag{A.5}$$

where $S_{ij} := e_i e_i^\top (\hat{y}_{ij} - e_i^\top Y e_j)^* + e_j e_i^\top (e_j^\top Y^* e_i) \in \mathbf{C}^{n\times n}$.

Proof.

(i) The complex voltage and complex current at bus $i \in \mathcal{V}$ can be expressed as $v_i = e_i^\top v$ and $I_i = e_i^\top I$, respectively. Therefore, the net complex power injection at bus i is equal to

$$s_i = e_i^\top v I^* e_i = e_i^\top v v^* Y^* e_i = \text{tr}(v v^* Y^* e_i e_i^\top) = v^* S_i v.$$

Equivalence in the last two equations follows from the cyclic property of the trace operator, which states that for any two matrices $A, B \in \mathbf{C}^{n \times n}$, $\text{tr}(AB) = \text{tr}(BA)$.

(ii) In a similar vein as part (i), we can write the complex power flow from bus i to bus j, as follows

$$\begin{aligned}
s_{ij} &= e_i^\top v I_{ij}^* \\
&= e_i^\top v \left((\hat{y}_{ij} - e_i^\top Y e_j) e_i^\top v + (e_i^\top Y e_j) e_j^\top v \right)^* \\
&= (\hat{y}_{ij} - e_i^\top Y e_j)^* e_i^\top v v^* e_i + (e_j^\top Y^* e_i) e_i^\top v v^* e_j \\
&= v^* \left(e_i e_i^\top (\hat{y}_{ij} - e_i^\top Y e_j)^* + e_j e_i^\top (e_j^\top Y^* e_i) \right) v \\
&= v^* S_{ij} v.
\end{aligned}$$

■

Remark 15 (Solving AC-Power Flow)**.** *There are several numerical methods in the literature, which are used to solve the power flow and power balance equations. Two popular popular methods are the Gauss-Seidel and Newton-Raphson. We refer the reader to [9] for a detailed discussion on such numerical methods.*

A.2 Linear Approximations of AC Power Flow Equations

A.2.1 DC Power Flow

In section A.1, we have seen that the AC power flow equations are quadratic in the bus voltage phasors, therefore they are nonlinear. This section provides a description of a linear approximation of these equations, known as DC power flow. The four assumptions governing this approximation are

(i) *Fixed voltage magnitudes:* All voltage magnitudes are equal to one per unit, i.e.,

$$|v_i| = 1, \qquad i \in \mathcal{V}.$$

(ii) *Small-angle approximation:* The voltage angle differences between neighboring nodes are small. Thus, $\sin(\theta_i - \theta_j) \approx \theta_i - \theta_j$ and $\cos(\theta_i - \theta_j) \approx 1$ for all $(i, j) \in \mathcal{E}$, where $\theta_i := \measuredangle v_i$.

(iii) *Lossless transmission lines:* This assumption implies that the resistance of the line connecting any two neighboring buses $(i, j) \in \mathcal{E}$ is equal to zero. Thus, the corresponding line admittance y_{ij} is purely imaginary and equal to

$$y_{ij} = \frac{r_{ij}}{r_{ij}^2 + x_{ij}^2} - \mathbf{i}\frac{x_{ij}}{r_{ij}^2 + x_{ij}^2} = -\mathbf{i}\frac{1}{x_{ij}},$$

where $r_{ij}, x_{ij} \in \mathbf{R}$ denote the resistance and reactance of the line form bus i to bus j, respectively. This assumption implies that the real part of the bus admittance matrix (A.1) is zero.

(iv) The shunt admittance $\hat{y}_i$ at each bus $i \in \mathcal{V}$ is equal to zero.

As one can readily verify, the above assumptions imply that the reactive power flows are equal to zero. Thus the complex power flow from bus i to bus j is purely real (active power) and equal to

$$p_{ij} := \mathrm{Re}\{s_{ij}\} = \frac{1}{x_{ij}}(\theta_i - \theta_j).$$

Moreover, the net power injected at bus i is given by

$$p_i := \mathrm{Re}\{s_i\} = B\theta,$$

where the matrix $B \in \mathbf{R}^{n \times n}$ is given by

$$[B]_{ij} := \begin{cases} -\dfrac{1}{x_{ij}}, & (i,j) \in \mathcal{E}, \\ \displaystyle\sum_{j:(i,j)\in\mathcal{E}} \frac{1}{x_{ij}}, & i = j, \\ 0, & \text{otherwise.} \end{cases} \tag{A.6}$$

We remark that B is a diagonally dominant matrix and thus positive semidefinite. We have the following Lemma which describes the DC power flow equations, which are linear in the bus voltage angles.

Lemma A.2.1 (DC Power Flow). *For a given power network on n buses, let $\theta \in \mathbf{R}^n$ be a vector of bus voltage angles and $p := \mathrm{Re}\{s\}$ be the vector of real power injections. Under assumptions ((i))-((iv)), the AC power flow equations (A.3) can be expressed as*

$$p = B\theta, \tag{A.7}$$

where $B \in \mathbf{R}^{n \times n}$ is defined in (A.6). Moreover, the power flow from bus i to bus j is given by

$$p_{ij} = e_i^\top B D_{ij}\theta, \qquad (i,j) \in \mathcal{E}, \tag{A.8}$$

where the matrix $D_{ij} \in \mathbf{R}^{n \times n}$ is defined by $D_{ij} := e_j(e_j - e_i)^\top$.

A.3 Concise Reformulation of RAC-OPF

Define matrices $\Phi_i, \Psi_i \in \mathbf{H}^n$, for all $i \in \mathcal{V}$, and a matrix $E \in \mathbf{R}^{4n\times 2n}$ as follows:

$$\Phi_i := \frac{S_i + S_i^*}{2}, \quad \Psi_i := \frac{S_i - S_i^*}{\mathbf{j}2}, \quad E := \begin{bmatrix} I_{2n} & -I_{2n} \end{bmatrix}^*.$$

In addition, let $f \in \mathbf{R}^{4n}$ be a vector given by

$$f := \left[\mathrm{Re}\{g^{\max}\}^* \;\; \mathrm{Im}\{g^{\max}\}^* \;\; -\mathrm{Re}\{g^{\min}\}^* \;\; -\mathrm{Im}\{g^{\min}\}^*\right]^*.$$

The RAC-OPF problem (5.8) can be reformulated as follows:

$$\begin{aligned}
\text{minimize} \quad & \mathbb{E}\left[\sum_{i=1}^{n} \alpha_i v(\boldsymbol{\xi})^* \Phi_i v(\boldsymbol{\xi})\right] + \sum_{i=1}^{n} \alpha_i \mathrm{Re}\{d_i\} \\
\text{subject to} \quad & x \in \mathbf{R}^{2n}, \quad v \in \mathbf{L}^2_{k,n} \\
& \left.\begin{array}{ll}
v(\xi)^* \Phi_i v(\xi) \le \mathrm{Re}\{e_i^* \overline{G} - d_i e_1^*\}^* \xi, & i \in \mathcal{V} \\
v(\xi)^* (-\Phi_i) v(\xi) \le \mathrm{Re}\{d_i e_1^* - e_i^* \underline{G}\}^* \xi, & i \in \mathcal{V} \\
v(\xi)^* \Psi_i v(\xi) \le \mathrm{Im}\{e_i^* \overline{G} - d_i e_1^*\}^* \xi, & i \in \mathcal{V} \\
v(\xi)^* (-\Psi_i) v(\xi) \le \mathrm{Im}\{d_i e_1^* - e_i^* \underline{G}\}^* \xi, & i \in \mathcal{V} \\
v(\xi)^* \Phi_i v(\xi) - e_i^* x \le \mathrm{Re}\{r_i^{\max} - d_i\} e_1^* \xi, & i \in \mathcal{V} \\
v(\xi)^* (-\Phi_i) v(\xi) + e_i^* x \le \mathrm{Re}\{d_i - r_i^{\min}\} e_1^* \xi, & i \in \mathcal{V} \\
v(\xi)^* \Psi_i v(\xi) - e_{n+i}^* x \le \mathrm{Im}\{r_i^{\max} - d_i\} e_1^* \xi, & i \in \mathcal{V} \\
v(\xi)^* (-\Psi_i) v(\xi) + e_{n+i}^* x \le \mathrm{Im}\{d_i - r_i^{\min}\} e_1^* \xi, & i \in \mathcal{V} \\
v(\xi)^* e_i e_i^* v(\xi) \le (v_i^{\max})^2 e_1^* \xi, & i \in \mathcal{V} \\
v(\xi)^* (-e_i e_i^*) v(\xi) \le -(v_i^{\min})^2 e_1^* \xi, & i \in \mathcal{V} \\
v(\xi)^* P_{ij} v(\xi) \le \ell_{ij}^{\max} e_1^* \xi, & (i,j) \in \mathcal{E} \\
v(\xi)^* (-P_{ij}) v(\xi) \le \ell_{ij}^{\max} e_1^* \xi, & (i,j) \in \mathcal{E}
\end{array}\right\} \; \forall\, \xi \in \Xi \\
& Ex \le f.
\end{aligned}$$

A.4 LMI Reformulation of Convex Inner Approximation $\mathcal{P}_{\mathrm{III}}(V_0)$

For $i = 0, 1, \ldots, m$, let $B_i = (A_i^+)^{1/2}$ and $N = M^{1/2}$ be the square roots of A_i^+ and M, respectively. These matrices are guaranteed to exist since both A_i^+ and M are positive semidefinite. In our reformulation, it will be convenient to write

$$\mathrm{tr}(MV^*A_0^+V) = \mathrm{tr}\left((B_0VN)^*(B_0VN)\right) = \mathrm{vec}(B_0VN)^*\mathrm{vec}(B_0VN),$$

where $\mathrm{vec}(\cdot)$ denotes the linear operator vectorizing matrices by stacking their columns. Let

$$L_i(V, V_0) = H_i(V, V_0) - V^*A_i^+V = V_0^*A_i^-V + V^*A_i^-V_0 - V_0^*A_i^-V_0$$

denote the part of $H_i(V, V_0)$ that depends affinely on V. Applying the Schur complement formula to the matrix inequalities in $\mathcal{P}_{\mathrm{III}}(V_0)$, we obtain the following equivalent reformulation of $\mathcal{P}_{\mathrm{III}}(V_0)$ as a semidefinite program:

$$\begin{array}{ll}
\text{minimize} & t + \mathrm{tr}(ML_0(V, V_0)) \\
\text{subject to} & x \in \mathbf{R}^{2n},\ V \in \mathbf{C}^{n\times k},\ \Lambda \in \mathbf{R}^{m\times \ell},\ t \in \mathbf{R} \\
& \begin{bmatrix} -I_n & B_iV \\ V^*B_i^* & L_i(V, V_0) - C_i + (b_i^*x)e_1e_1^* - \displaystyle\sum_{j=1}^{\ell}[\Lambda]_{ij}W_j \end{bmatrix} \preceq 0, \quad i = 1, \ldots, m \\
& \begin{bmatrix} -tI_{nk} & \mathrm{vec}(B_0VN) \\ \mathrm{vec}(B_0VN)^* & -1 \end{bmatrix} \preceq 0, \\
& Ex \leq f, \\
& \Lambda \leq 0.
\end{array}$$

APPENDIX B

CONVEX OPTIMIZATION

B.1 Karush Kuhn Tucker (KKT) Conditions for Local Optimality

In this section, we review the first-order necessary conditions for local optimality developed by Karush (1939) and Kuhn and Tucker (1951). Consider the constrained nonlinear optimization problem

$$\begin{aligned} \underset{x \in \mathbf{R}^n}{\text{minimize}} \qquad & f(x) \\ \text{subject to} \qquad & g(x) \leq 0, \end{aligned} \tag{B.1}$$

where $f : \mathbf{R}^n \to \mathbf{R}$ and $g : \mathbf{R}^n \to \mathbf{R}^m$ are assumed to be continuously differentiable. Therefore, for every $x \in \mathbf{R}^n$ and for each $i = 1, \ldots, m$,

$$\nabla f(x) \coloneqq \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} & \frac{\partial f(x)}{\partial x_2} & \cdots & \frac{\partial f(x)}{\partial x_n} \end{bmatrix}^\top,$$

$$\nabla g_i(x) \coloneqq \begin{bmatrix} \frac{\partial g_i(x)}{\partial x_1} & \frac{\partial g_i(x)}{\partial x_2} & \cdots & \frac{\partial g_i(x)}{\partial x_n} \end{bmatrix}^\top,$$

exist and vary continuously with x. We define $\nabla g(x) \coloneqq [\nabla g_1(x), \ldots, \nabla g_m(x)] \in \mathbf{R}^{n \times m}$.

Definition B.1.1 (Local minimum)**.** *We say that $\overline{x}$ is a local minimizer of problem* (B.1) *if it is feasible and if no sufficiently close feasible point has a better objective value. That is, for some $\varepsilon > 0$, $\|x - \overline{x}\| \leq \varepsilon$ and $g(x) \leq 0$ imply that $f(x) \geq f(\overline{x})$.*

Definition B.1.2 (MFCQ)**.** *Given $\overline{x} \in \mathbf{R}^n$ with $g(\overline{x}) \leq 0$, let*

$$I(\overline{x}) \coloneqq \{i \mid 1 \leq i \leq m,\ g_i(\overline{x}) = 0\}$$

denote the set of active constraint indices at $\overline{x}$. The Mangasarian-Fromovitz constraint qualification (MFCQ) holds at $\overline{x}$ if there is no nontrivial nonnegative linear dependence among the $\nabla g_i(\overline{x}),\ i \in I(\overline{x})$.

An important consequence is that if MFCQ holds at a feasible point $\overline{x}$, then there is some $d \in \mathbf{R}^n$, which satisfies $\nabla g_i(x)^\top d < 0$, for all $i \in I(\overline{x})$. Moreover, such a d is a minimizer of the direction-finding subproblem

$$\begin{aligned} \underset{d\in\mathbf{R}^n}{\text{minimize}} \quad & \nabla f(\overline{x})^\top d \\ \text{subject to} \quad & \nabla g(\overline{x})^\top d \leq -g(\overline{x}), \end{aligned} \tag{B.2}$$

which, modulus the constant term $f(\overline{x})$, aims to minimize the first-order Taylor approximation of $f(\overline{x}+d)$ subject to requiring that the first-order Taylor approximations of all $g_i(\overline{x}+d)$, $i = 1,\ldots,m$, are nonpositive. Problem (B.2) is a linear programming problem, a linearization of (B.1). The KKT necessary conditions for local optimality, are described by the following Theorem.

Theorem B.1.3 (KKT Conditions). *Suppose that $\overline{x}$ is a local minimizer for* (B.1) *and that the MFCQ holds at $\overline{x}$. Then, there are Lagrange multipliers $\overline{u} \in \mathbf{R}^m$ such that the following conditions hold:*

1. *Stationarity:* $\nabla f(\overline{x}) + \nabla g(\overline{x})\overline{u} = 0$
2. *Primal Feasibility:* $g(\overline{x}) \leq 0$
3. *Dual Feasibility:* $\overline{u} \geq 0$
4. *Complementary Slackness:* $\overline{u}_i g_i(\overline{x}) = 0, \ \forall\, i = 1,\ldots,m.$

The first condition says that the gradient of the objective function plus a nonnegative combination of the active constraints gives zero. The second constraint ensures that $\overline{x}$ satisfies the constraints. Finally, the last constraint ensures that $u_i > 0$, only if i is an index of an active constraint.

Algorithm 1: Alternating Direction Method of Multipliers

Given an initial condition $(x^0, z_1^0, \ldots, z_{m+\ell}^0, u^0)$, a stopping tolerance $\varepsilon > 0$, and maximum number of iterations $\overline{k}$
Initialize $k = 0$
Repeat

1. *Compute.* $x^{k+1} = \underset{z_i}{\text{argmin}}\ L_\rho(x, z^k, u^k) = \underset{x}{\text{argmin}}\ g(x) + \rho \sum_{i=1}^m \|z_i - x + u_i\|_2^2$
2. For each $i = 1, \ldots, m$:
   (i) *Compute.* $z_i^{k+1} = \underset{z_i}{\text{argmin}}\ L_\rho(x_{k+1}, z_i, u^k) = \underset{z_i}{\text{argmin}}\ f_i(z_i) + \rho\|z_i - x^{k+1} + u_i^k\|_2^2$
   (ii) *Compute.* $u_i^{k+1} = u_i^k + z_i^{k+1} - x^{k+1}$
3. *Update.* $k = k + 1$

Until $|x^{k+1} - x^k| < \varepsilon,$ and $|z^i - x_k| < \varepsilon,$ for all $i = 1, \ldots, m + \ell$
Output $(x^k, z_1^k, \ldots, z_{m+\ell}^k, u^k)$

Table B.1: General Consensus Alternating Direction Method of Multipliers Algorithm.

B.2 Alternating Direction Method of Multipliers

Consider the following optimization problem

$$\underset{x \in \mathbf{R}^n}{\text{minimize}}\ \ g(x) + \sum_{i=1}^m f_i(x) \tag{B.3}$$

in which the objective is to minimize the sum of cost functions $f_1, \ldots, f_m : \mathbf{R}^n \to \mathbf{R}$ plus an additional regularization term, which is described by the function $g : \mathbf{R}^n \to \mathbf{R}$. Clearly, the above unconstrained optimization problem can be expressed in the following form

$$\begin{aligned} \underset{x, z_1, \ldots, z_m \in \mathbf{R}^n}{\text{minimize}} \quad & g(x) + \sum_{i=1}^m f_i(z_i) \\ \text{subject to} \quad & z_i = x, \quad i = 1, \ldots, m, \end{aligned} \tag{B.4}$$

where we have introduced m additional decision variables $z_1, \ldots, z_m \in \mathbf{R}^n$ to decouple the objective function. These variables are called *consensus variables* and are coupled by equality constraints in problem (B.4). The augmented Lagrangian function associated with the above optimization problem is given by

$$L_\rho(x, z_1, \dots, z_m, y) = g(x) + \sum_{i=1}^{m} f_i(z_i) + \sum_{i=1}^{m} u_i^\top (z_i - x) + \sum_{i=1}^{m} \rho \|z_i - x\|_2^2.$$

where, the variable u_i, $i = 1, \dots, m$ is a dual variable corresponding to the equality constraint $z_i = x$. The consensus ADMM algoritm for problem (B.4) is described in Table B.1 where the superscript $k = 0, 1, \dots$ denotes the k^{th} iterate. ‘

B.3 Conic Programming

Let $\mathbf{E}$ be a real n-dimensional Euclidean space with inner product $\langle \cdot, \cdot \rangle$. A subset $\mathcal{K} \subseteq \mathbf{E}$ is a convex cone if it contains zero and $ax + by \in \mathcal{K}$ for any nonnegative scalars a, b and any two points $x, y \in \mathcal{K}$. The *dual* cone of $\mathcal{K}$ is given by $\mathcal{K}^* = \{y \in \mathbf{R}^n \mid \langle x, y \rangle \geq 0, \text{ for all } x \in \mathcal{K}\}$. We say that $\mathcal{K}$ is self-dual if $\mathcal{K} = \mathcal{K}^*$. A cone $\mathcal{K}$ is said to be *pointed* if $\mathcal{K} \cap (-\mathcal{K}) = \{0\}$. It is said to be *proper*, if it is closed, convex, pointed, and with a non-empty interior. The dual cone of a proper cone is also a proper cone.

Let $\mathcal{K} \subseteq \mathbf{E}$ be a closed convex cone, $A : \mathbf{E} \to \mathbf{R}^m$ a linear operator [1] and $c \in \mathbf{E}$ and $b \in \mathbf{R}^m$ be two vectors. In its most general form, a conic linear program and its dual are given by

$$\begin{array}{lll} \underset{x \in \mathbf{E}}{\text{minimize}} & \langle c, x \rangle \\ \text{subject to} & Ax = b \\ & x \in \mathcal{K} \end{array} \qquad \begin{array}{ll} \underset{y \in \mathbf{R}^m,\ z \in \mathbf{E}}{\text{maximize}} & \langle b, y \rangle \\ \text{subject to} & z + A^* y = c \\ & z \in \mathcal{K}^*. \end{array} \tag{B.5}$$

This pair of primal-dual optimization problems includes the following families of problems.

[1] Without loss of generality, we can assume that $A : \mathbf{E} \to \mathbf{R}^m$ is surjective.

(i) *Linear Programs,* when $\mathbf{E} = \mathbf{R}^n,\ \langle x, y\rangle = x^\top y$ and $\mathcal{K}$ is the nonnegative orthant, i.e.,

$$\mathcal{K} = \mathbf{R}^n_+ := \{x \in \mathbf{R}^n \mid x \geq 0\}.$$

(ii) *Second-Order Cone Programs:* when $\mathbf{E} = \mathbf{R}^n$, and $\langle x, y\rangle = x^\top y$ and and $\mathcal{K}$ is the Lorentz (second-order) cone, i.e.,

$$\mathcal{K} = \mathbf{L}^n := \{(t, x) \in \mathbf{R} \times \mathbf{R}^{n-1} \mid \|x\| \leq t\}.$$

(iii) *Positive Semidefinite Programs:* when $\mathbf{E} = \mathbf{S}^n$, $\langle X, Y\rangle = \mathrm{tr}(XY)$, and $Kcal$ is the positive semidefinite cone, i.e.,

$$\mathcal{K} = \mathbf{S}^n_+ = \{X \in \mathbf{S}^n \mid y^\top X y \geq 0, \text{ for all } y \in \mathbf{R}^n\}.$$

Among the three cones mentioned above, only the nonnegative orthant is polyhedral, but all three of them are self-dual.

B.4 Constraint Nondegeneracy for Semidefinite Programs

Consider the semidefinite program

$$\begin{aligned} \underset{X \in \mathbf{H}^n}{\text{minimize}} \quad & \mathrm{tr}(A_0 X) \\ \text{subject to} \quad & \mathrm{tr}(A_i X) \leq b_i, \qquad i = 1, \ldots, m, \\ & \mathrm{tr}(A_i X) = b_i, \qquad i = m+1, \ldots, m+\ell, \\ & X \succeq 0. \end{aligned} \tag{B.6}$$

where the matrices $A_0, A_1, \ldots, A_{m+\ell} \in \mathbf{H}^n$ and the scalars $b_1, \ldots, b_{m+\ell} \in \mathbf{R}$ are the given problem data. The dual problem to (B.6) is given by

$$
\begin{aligned}
\text{minimize} \quad & b^\top y \\
\text{subject to} \quad & y \in \mathbf{R}^m, \\
& A_0 - \sum_{i=1}^{m+\ell} y_i A_i \in \mathbf{H}^n_+, \\
& y_i \geq 0, \quad i = 1, \ldots, m.
\end{aligned}
$$

Given a primal feasible solution $X \in \mathbf{H}^n$, let us denote by

$$
\mathcal{I}(X) := \{i \mid 1 \leq i \leq m, \ \mathrm{tr}(A_i X) = b_i\}.
$$

the indices of the inequality constraints that are active at X. In what follows we define *nondegeneracy* of primal and dual feasible solutions.

Definition 1 (Primal Nondegeneracy). *Let X be a primal feasible solution and suppose that* $\mathrm{rank}(X) = r$. *Let $X = Q\Lambda Q^*$ be an eigenvalue decomposition of X, where* $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_r, 0, \ldots, 0) \in \mathbf{R}^{n\times n}$ *and $Q \in \mathbf{C}^{n\times n}$. Partition Q as $Q = [Q_1, \ Q_2]$, where $Q_1 \in \mathbf{C}^{n\times r}$ and $Q_2 \in \mathbf{C}^{n\times(n-r)}$, and define the matrices*

$$
B_k^p := \begin{bmatrix} Q_1^* A_k Q_1 & Q_1^* A_k Q_2 \\ Q_2^* A_k Q_1 & 0 \end{bmatrix}, \quad k \in I(X) \cup \{m+1, \ldots, m+\ell\}.
$$

Then X is primal nondegenerate*, if the matrices $B_k^p, \ k \in I(X) \cup \{m+1, \ldots, m+\ell\}$ are linearly independent in $\mathbf{H}^n$.* ■

Definition 2 (Dual Nondegeneracy). *Let (y, Z) be a dual feasible point and suppose that* $\mathrm{rank}(Z) = s$. *Let $Z = P\Sigma P^*$ be an eigenvalue decomposition of Z, where* $\Sigma = \mathrm{diag}(0, \ldots, 0, \sigma_{n-s+1}, \ldots, \sigma_n) \in \mathbf{R}^{n\times n}$ *and $P \in \mathbf{C}^{n\times n}$. Partition P as $P = [P_1 \ \ P_2]$ where $P_1 \in \mathbf{C}^{n\times(n-s)}$ and $P_2 \in \mathbf{C}^{n\times s}$, and define the matrices*

$$
B_k^d := P_1^* A_k P_1, \quad k \in I(X) \cup \{m+1, \ldots, m+\ell\}.
$$

Then (y, Z) *is* dual nondegenerate*, if and only if the matrices* $B_k^d,\ k \in I(X) \cup \{m+1, \ldots, m+\ell\}$ *span* $\mathbf{H}^{n-s}$. ■

Note that if (X, y, Z) is a primal-dual optimal solution pair that satisfies strict complementarity (i.e., $\operatorname{rank}(X) + \operatorname{rank}(Z) = n$), then $Q_1 = P_1$ and $Q_2 = P_2$.

Remark 16 (Transversality)**.** *Primal nondegeneracy has the following geometric interpretation. Let* $\mathcal{M}_r := \{X \in \mathbf{H}^n \mid \operatorname{rank}(X) = r\}$ *be the set of Hermitian matrices of order* n *that have rank* r*. A primal feasible point* $X \in \mathcal{M}_r$ *is nondegenerate if and only if the orthogonal complement of the subspace spanned by the matrices* $A_k,\ k \in,$ *intersects the tangent space to* $\mathcal{M}_r$ *at* X *transversally. See [1, 73] for a definition of transversality.*

Primal and dual nondegeneracy is related to uniqueness of optimal solutions to semidefinite programs. In particular, we have the following result from Alizadeh et al. [1].

Theorem 1 (Uniqueness of Optimal Solutions)**.** *Let* y *be dual nondegenerate and optimal. Then, there exists a unique primal optimal solution. Similarly, if* X *is primal nondegenerate and optimal, then there exists a unique dual optimal solution.*

B.5 The S-Procedure

The S-procedure provides a sufficient condition for proving set containments involving quadratic functions by verifying the feasiblity of a matrix inequality which is linear in the data defining said quadratic functions. We have the following Proposition

Proposition B.5.1 (S-Procedure)**.** *Let* $f_0, \ldots, f_m$ *be quadratic functions of* $x \in \mathbf{R}^n$*, i.e.,*

$$f_i(x) := x^* A_i x + 2b_i^* x + c_i, \qquad i = 0, \ldots, m,$$

where $A_i \in \mathbf{S}^n$, $b_i \in \mathbf{R}^n$, and $c_i \in \mathbf{R}$, for all $i = 0, \ldots, m$. The condition

$$f_0(x) \geq 0, \text{ for all } x \text{ such that } f_i(x) \leq 0, \ i = 1, \ldots, m,$$

holds if there exists scalars $\tau_i \geq 0$, $i = 1, \ldots, m$ such that

$$\begin{bmatrix} A_0 & b_0 \\ b_0^\top & c_0 \end{bmatrix} + \sum_{i=1}^{p} \begin{bmatrix} A_i & b_i \\ b_i^\top & c_i \end{bmatrix} \succeq 0.$$

Moreover, if $m = 1$, the converse holds provided that there is some $\overline{x} \in \mathbf{R}^n$ such that $f_1(\overline{x}) > 0$.

B.6 The Schur Complement

Let X be a Hermitian $n \times n$ matrix. Suppose that X is decomposed into a 2×2 block form

$$X := \begin{bmatrix} A & B \\ B^* & C \end{bmatrix}, \tag{B.7}$$

where A is a $p \times p$ Hermitian matrix and C is a $q \times q$ Hermitian matrix such that $n = p + q$. It follows that B is a complex $p \times q$ matrix. If A is nonsingular, then the Schur complement of a block Hermitian matrix X of the form (B.7) is given by

$$S := C - B^\top A^{-1} B.$$

It provides a condition guaranteeing that X is positive (semi)definite, which depends on the positive (semi)definitness of the block matrices A, B, C. More precisely, we have the following Lemma.

Lemma B.6.1. *Let $X \in \mathbf{H}^n$ be a Hermitian matrix of the form (B.7). If A is invertible, then the following properties hold:*

(i) $X \succ 0$*, if and only if* $A \succ 0$ *and* $C - BA^{-1}B^* \succ 0$*,*

(ii) *If* $A \succ 0$*, then* $X \succeq 0$ *if and only if* $C - BA^{-1}B^* \succeq 0$

B.7 Quadratically Constrained Quadratic Programs

A quadratically constrained quadratic program (QCQP) is a nonlinear optimization problem that can be expressed in the form

$$\begin{aligned} \underset{x \in \mathbf{C}^n}{\text{minimize}} \quad & x^* A_0 x \\ \text{subject to} \quad & x^* A_k x \leq b_k, \quad \text{for all } k = 1, \ldots, m. \end{aligned} \tag{B.8}$$

where the scalars $b_1, \ldots, b_m \in \mathbf{R}$ and the matrices $A_0, A_1, \ldots, A_m \in \mathbf{H}^n$ are the given problem data. If $A_i \succeq 0$ for all $i = 0, 1, \ldots, m$, then (B.9) is equivalent to the second order cone program

$$\begin{aligned} \underset{x \in \mathbf{C}^n}{\text{minimize}} \quad & \|L_0 x\|_2^2 \\ \text{subject to} \quad & \|L_k x\|_2^2 \leq b_k, \quad \text{for all } k = 1, \ldots, m, \end{aligned} \tag{B.9}$$

where $A_k = L_k^* L_k, \ k = 0, \ldots, m$ is the Cholesky factorization of A_k. However, we do not make this assumption here. The Lagrangian of (B.9) is given by

$$L(x, \lambda) = x^* \left(A_0 + \sum_{i=1}^{m} \lambda_i A_i \right) x - b^* \lambda.$$

The dual problem of (B.9) is a semidefinite program and it is given by

$$\begin{aligned} \underset{\lambda \in \mathbf{R}}{\text{maximize}} \quad & - b^* \lambda \\ \text{subject to} \quad & A_0 + \sum_{i=1}^{m} \lambda_i A_i \succeq 0, \\ & \lambda \geq 0. \end{aligned} \tag{B.10}$$

And the Lagrangian bidual of the QCQP (B.9) (i.e., the dual problem of the semidefinite program (B.10)) is equal to the semidefinite relaxation of the QCQP. More precisely, the dual problem of (B.10) is given by

$$\begin{aligned} \underset{X \in \mathbf{H}^n}{\text{minimize}} \quad & \operatorname{tr}(A_0 X) \\ \text{subject to} \quad & \operatorname{tr}(A_i X) \leq b_i, \qquad i = 1, \ldots, m \\ & X \succeq 0. \end{aligned} \tag{B.11}$$

The following Theorem from [5] gives rise .

Theorem B.7.1. *Suppose that there exists a positive semidefinite matrix $X \in \mathbf{H}_+^n$ to the system of equations*

$$\operatorname{tr}(A_k X) = b_k, \quad k = 1, \ldots, m.$$

Then there is a positive semidefinite matrix $\overline{X} \in \mathbf{H}_+^n$ to the above system such that

$$\operatorname{rank}(\overline{X}) \leq -1 + \sqrt{1+m}.$$

A direct implication of the above Theorem is that nonconvex QCQPs with at most two constraints can be solved in polynomial time through their semidefinite relaxation. We have the following Corollary.

Corollary B.7.2. *Let $m = 2$ and suppose that $X^\star \in \mathbf{H}_+^n$ is an optimal solution of the semidefinite relaxation of the nonconvex QCQP (B.9). Consider the system of linear equations*

$$\operatorname{tr}(A_k X) = \operatorname{tr}(A_k X^\star), \qquad k = 0, \ldots, m.$$

Then, there is a positive semidefinite matrix $\overline{X} \in \mathbf{H}_+^n$ to the above system of equations such that

$$\operatorname{rank}(\overline{X}) \leq 1.$$

B.8 Polyhedral Approximations of Semidefinite Programs

In this section, we consider a recursive method for constructing outer and inner polyhedral approximations to semidefinite programs. Consider the semidefinite program

$$\begin{aligned} \underset{x \in \mathbf{R}^m}{\text{minimize}} \quad & c^\top x \\ \text{subject to} \quad & A_0 + \sum_{i=1}^{m} x_i A_i \in \mathbf{S}_+^n, \end{aligned} \tag{B.12}$$

where the vector $c \in \mathbf{R}^m$ and the symmetric matrices $A_0, \ldots, A_i \in \mathbf{S}^n$ are the given problem data. Let

$$\mathbf{P} := \bigcap_{j=1}^{p} \{X \in \mathbf{S}^n \mid \text{tr}(Z_j X) \geq 0\}, \tag{B.13}$$

be an arbitrary polyhedral cone, described by matrices $Z_1, \ldots, Z_p \in \mathbf{S}_+^n$. It follows by the self-duality of the positive semidefinite cone that that $\mathbf{P} \supseteq \mathbf{S}_+^n$. Henceforth, we refer to $\mathbf{P}$ as the *outer polyhedral cone*. Therefore, the linear program

$$\begin{aligned} \underset{x \in \mathbf{R}^m}{\text{minimize}} \quad & c^\top x \\ \text{subject to} \quad & A_0 + \sum_{i=1}^{m} x_i A_i \in \mathbf{P}, \end{aligned} \tag{B.14}$$

yields an outer approximation to the semidefinite program (B.12) and its optimal value stands as a lower bound to the optimal value of (B.12). Consider now the linear program

$$\begin{aligned} \underset{x \in \mathbf{R}^m}{\text{minimize}} \quad & c^\top x \\ \text{subject to} \quad & A_0 + \sum_{i=1}^{m} x_i A_i \in \mathbf{P}^*, \end{aligned} \tag{B.15}$$

where $\mathbf{P}^*$ denotes the dual cone of $\mathbf{P}$. It is given by

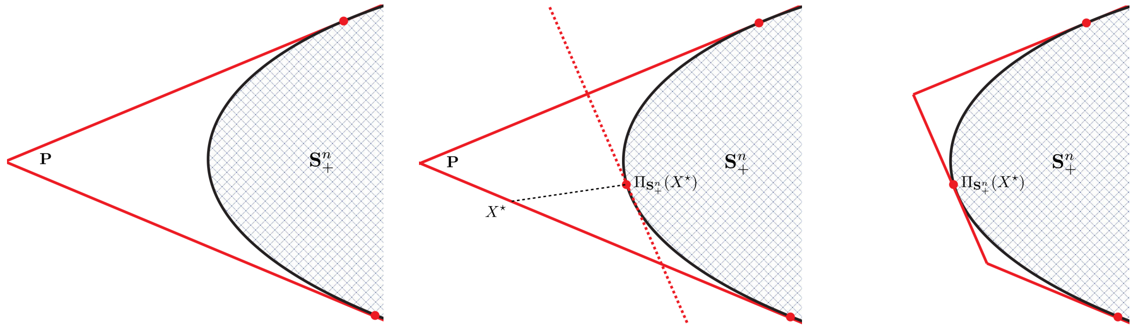
$$\mathbf{P}^* = \text{cone}\{Z_1, \ldots, Z_p\} = \left\{ \sum_{j=1}^{p} y_j Z_j \;\middle|\; y \geq 0 \right\}. \tag{B.16}$$

Since for any two convex cones K_1, K_2, which satisfy $K_1 \subseteq K_2$, it holds that $K_1^* \subseteq K_2^*$, we obtain by the self-duality of the positive semidefinite cone that the linear program (B.15) is an inner approximation to the semidefinite program (B.12). Its optimal value stands as an upper bound to the optimal value of (B.12).

B.8.1 Recursive Polyhedral Approximations of Semidefinite Programs

The effectiveness of the polyhedral approximations introduced in the previous section depends critically on the choice of the outer polyhedral cone $\mathbf{P}$, and for any given problem it is unclear what the best choice for $\mathbf{P}$ is. A naive approach might entail the construction of a hierarchy of polyhedral cones via a uniform discretization of the boundary of the cross polytope [46]. For high levels in the hierarchy, however, this approach could yield computational inefficiencies due to the large number of half-spaces defining the resulting polyhedral cones. In turn, this gives rise to linear programs (B.14) and (B.15) with a large number of variables and constraints.

In this section, we explore polyhedral approximations of the positive semidefinite cone that are adaptively guided by the objective function. More precisely, starting with a coarse outer polyhedral cone (i.e., the cone of symmetric matrices with nonnegative diagonal entries), we prescribe a recursive method to refine this cone, which uses the solution at the current iteration step. Qualitatively, at each iteration step, we project the optimal solution onto the positive semidefinite cone and refine the outer polyhedral cone by intersecting it with the half-space corresponding to the supporting hyperplane at said projection point. The resulting feasible set of the outer approximation of the semidefinite program (B.12) over the refined cone is shown to exclude the optimal solution at the previous iteration step.



(a) (b) (c)

Figure B.1: Figure B.1(a) shows the positive semidefinite cone and the outer polyhedral cone $\mathbf{P}$, which is described by two half-spaces. In Figure B.1(b), we visualize (illustrated by the dotted line) the supporting hyperplane to $\mathbf{S}_+^n$ at $\Pi_{\mathbf{S}_+^n}(X^\star)$. In Figure B.1(c), we visualize the refined outer polyhedral cone obtained by intersecting $\mathbf{P}$ with the half-spaces corresponding to the supporting hyperplane.

Let $x \in \mathbf{R}^m$ be a feasible solution to the outer approximation (B.14). A cutting plane for problem (B.14) is defined to be a hyperplane characterized by a matrix $Z \in \mathbf{S}_+^n$ and passing through the origin such that

$$A_0 + \sum_{i=1}^{m} x_i A_i \in \{X \in \mathbf{S}^n \mid \mathrm{tr}(ZX) < 0\}. \tag{B.17}$$

An important consequence of (B.17) is that the outer approximation (B.14) obtained by approximating the positive semidefinite cone by $\mathbf{P} \cap \{X \in \mathbf{S}^n \mid \mathrm{tr}(ZX) \geq 0\}$ will not contain x in its feasible set.

Given an optimal solution $x^\star \in \mathbf{R}^m$ to the outer program (B.14), let us define, for the sake of simplicity, the matrix

$$X^\star := A_0 + \sum_{i=1}^{m} x_i^\star A_i. \tag{B.18}$$

The following Proposition shows that if $X^\star$ is not positive semidefinite, then the supporting hyperplane to $\mathbf{S}_+^n$ at $\Pi_{\mathbf{S}_+^n}(X^\star)$ is a cutting plane for (B.14).

Proposition B.8.1. *Let $x^\star$ be a primal optimal solution to the outer approximation* (B.14)

Algorithm

Given an outer polyhedral cone $\mathbf{P}$ and a maximum number $\bar{t}$ of iterations
Initialize $t = 1$
Repeat

1. *Let.* $p =$ number of half-spaces defining $\mathbf{P}$
2. *Compute.* An optimal solution $x^\star$ to the outer program (B.14) and a corresponding matrix $X^\star$ defined according to (B.18).
3. *Update.*
   - $\mathbf{P} = \mathbf{P} \bigcap_{j=1}^{p} \{X \in \mathbf{S}^n \mid \mathrm{tr}\big((\Pi_{\mathbf{S}^n_+}(X^\star) - X^\star)X\big) \geq 0\}$
   - $t = t + 1$

Until $X^\star \in \mathbf{S}^n_+$, or $t = \bar{t}$
Output The optimal value of the outer program (B.14) and the optimal value of the inner program B.15

Table B.2: Cutting plane algorithm to improve polyhedral approximations to the semidefinite program (B.12)

and define a matrix $X^\star$ according to (B.18) . *If $X^\star \notin \mathbf{S}^n_+$, then the hyperplane*

$$\{X \in \mathbf{S}^n \mid \mathrm{tr}\big((\Pi_{\mathbf{S}^n_+}(X^\star) - X^\star)X\big) = 0\}$$

is a cutting plane for (B.14).

Proof. We must show that

$$\mathrm{tr}((\Pi_{\mathbf{S}^n_+}(X^\star) - X^\star)X^\star) < 0$$

Let $X^\star = \sum_{i=1}^{n} \lambda_i u_i u_i^\top$ be an eigenvalue decomposition of $X^\star$. It follows that

$$\Pi_{\mathbf{S}^n_+}(X^\star) - X^\star = -\sum_{i=1}^{n} \min\{0, \lambda_i\} u_i u_i^\top.$$

Since the vectors $u_1, \ldots, u_n$ are orthonormal, we must have that

$$\mathrm{tr}((\Pi_{\mathbf{S}^n_+}(X^\star) - X^\star)X^\star) = -\sum_{i=1}^{n}\sum_{j=1}^{n} \min\{0, \lambda_i\}\lambda_j \mathrm{tr}(u_i u_i^\top u_j u_j^\top) = -\sum_{i=1}^{n} \min\{0, \lambda_i\}\lambda_i < 0.$$

■

In Table B.2, we outline the steps of the proposed algorithm. Starting with a coarse outer polyhedral cone $\mathbf{P}$, the algorithm computes an optimal solution $x^\star$ to the linear program (B.14). A cutting plane for (B.14) is then computed according to Proposition B.8.1. The intersection of $\mathbf{P}$ with the corresponding half-space yields a refined outer polyhedral cone. The above procedure is repeated using the refined cone until all the optimal solution to (B.14) yields a positive semidefinite matrix $X^\star$ or until a maximum number of iterations is reached.

Remark 17. *The recursive algorithm presented in this section is similar in spirit to the cutting plane algorithm developed in [65] by Krishnan and Mitchell. As argued in a follow-up paper [54] by the second author, in practice the cutting plane algorithm needs to add a large number of hyperplanes to* $\mathbf{P}$. *In addition, the cutting plane method is more competitive than primal-dual interior point methods for semidefinite programs only for high dimensional semidefinite programs having a small number of constraints. Furthermore, cutting plane methods are not able to solve semidefinite programs as accurately as interior point methods in comparable time. As an alternative to cutting plane methods, Helmberg and Rendl developed a spectral bundle method for semidefinite programs [30], which demonstrates excellent computational advantages for problems that are inaccessible to interior point methods.*

BIBLIOGRAPHY

[1] Farid Alizadeh, Jean-Pierre A Haeberly, and Michael L Overton. Complementarity and nondegeneracy in semidefinite programming. *Mathematical Programming*, 77(1):111–128, 1997.

[2] O Alsac and B Stott. Optimal load flow with steady-state security. *Power Apparatus and Systems, IEEE Transactions on*, (3):745–751, 1974.

[3] X. Bai, L. Qu, and W. Qiao. Robust AC optimal power flow for power networks with wind power generation. *IEEE Transactions on Power Systems*, PP(99):1–2, 2015.

[4] Xiaoqing Bai, Hua Wei, Katsuki Fujisawa, and Yong Wang. Semidefinite programming for optimal power flow problems. *International Journal of Electrical Power & Energy Systems*, 30(6):383–392, 2008.

[5] Alexander Barvinok. *A course in convexity*, volume 54. AMS Bookstore, 2002.

[6] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust optimization*. Princeton University Press, 2009.

[7] Aharon Ben-Tal, Alexander Goryashko, Elana Guslitzer, and Arkadi Nemirovski. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2):351–376, 2004.

[8] Aharon Ben-Tal, Arkadi Nemirovski, and Cees Roos. Robust solutions of uncertain quadratic and conic-quadratic problems. *SIAM Journal on Optimization*, 13(2):535–560, 2002.

[9] Arthur R.. Bergen and Vijay Vittal. *Power systems analysis*. Prentice Hall, 2000.

[10] Daniel Bienstock, Michael Chertkov, and Sean Harnett. Chance-constrained optimal power flow: Risk-aware network control under uncertainty. *SIAM Review*, 56(3):461–495, 2014.

[11] Grigoriy Blekherman, Pablo A Parrilo, and Rekha R Thomas. *Semidefinite Optimization and Convex Algebraic Geometry*. Siam, 2013.

[12] Jonathan Borwein and Adrian S Lewis. *Convex analysis and nonlinear optimization: theory and examples*. Springer Science & Business Media, 2010.

[13] Subhonmesh Bose, Dennice F Gayme, K Mani Chandy, and Steven H Low. Quadratically constrained quadratic programs on acyclic graphs with application to power flow. *IEEE Transactions on Control of Network Systems*, 2(3):278–287, 2015.

[14] Subhonmesh Bose, Steven H Low, Thanchanok Teeraratkul, and Babak Hassibi. Equivalent relaxations of optimal power flow. *IEEE Transactions on Automatic Control*, 60(3):729–742, 2015.

[15] Stephen P Boyd, Laurent El Ghaoui, Eric Feron, and Venkataramanan Balakrishnan. *Linear matrix inequalities in system and control theory*, volume 15. SIAM, 1994.

[16] RC Burchett, HH Happ, and DR Vierath. Quadratically convergent optimal power flow. *IEEE Transactions on Power Apparatus and Systems*, (11):3267–3275, 1984.

[17] J Carpentier. Contribution a letude du dispatching economique. *Bulletin de la Societe Francaise des Electriciens*, 3(1):431–447, 1962.

[18] Anya Castillo and Richard P O'Neill. *Survey of Approaches to Solving the ACOPF*. FERC, 2013.

[19] CY Chung, Wei Yan, and Fang Liu. Decomposed predictor-corrector interior point method for dynamic optimal power flow. *IEEE Transactions on Power Systems*, 26(3):1030–1039, 2011.

[20] Yonina C Eldar, Aharon Ben-Tal, and Arkadi Nemirovski. Robust mean-squared error estimation in the presence of model uncertainties. *IEEE Transactions on Signal Processing*, 53(1):168–181, 2005.

[21] Maryam Fazel, Haitham Hindi, and Stephen P Boyd. A rank minimization heuristic with application to minimum order system approximation. In *American Control Conference, 2001. Proceedings of the 2001*, volume 6, pages 4734–4739. IEEE, 2001.

[22] Maryam Fazel, Haitham Hindi, and Stephen P Boyd. Log-det heuristic for matrix rank minimization with applications to hankel and euclidean distance matrices. In *American Control Conference, 2003. Proceedings of the 2003*, volume 3, pages 2156–2162. IEEE, 2003.

[23] Robert M Freund and Jorge R Vera. Some characterizations and properties of the distance to ill-posedness and the condition measure of a conic linear system. *Mathematical Programming*, 86(2):225–260, 1999.

[24] Alex B Gershman, Nicholas D Sidiropoulos, Shahram Shahbazpanahi, Mats Bengtsson, and Björn Ottersten. Convex optimization-based beamforming: From receive to transmit and network designs. *IEEE Signal Process. Mag*, 27(3):62–75, 2010.

[25] Chris Godsil and Gordon F Royle. *Algebraic graph theory*, volume 207. Springer Science & Business Media, 2013.

[26] Michel X Goemans and David P Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)*, 42(6):1115–1145, 1995.

[27] D Goldfarb and K Scheinberg. Interior point trajectories in semidefinite programming. *SIAM Journal on Optimization*, 8(4):871–886, 1998.

[28] Simai He, Zhi-Quan Luo, Jiawang Nie, and Shuzhong Zhang. Semidefinite relaxation bounds for indefinite homogeneous quadratic optimization. *SIAM Journal on Optimization*, 19(2):503–523, 2008.

[29] Christoph Helmberg. *Semidefinite programming for combinatorial optimization*. Citeseer, 2000.

[30] Christoph Helmberg and Franz Rendl. A spectral bundle method for semidefinite programming. *SIAM Journal on Optimization*, 10(3):673–696, 2000.

[31] Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge university press, 2012.

[32] Kejun Huang and Nicholas D Sidiropoulos. Consensus-admm for general quadratically constrained quadratic programming. *IEEE Transactions on Signal Processing*, 64(20):5297–5310.

[33] Edison Electric Institute. Key facts about the electric power industry.

[34] Rabih A Jabr. Radial distribution load flow using conic programming. *IEEE transactions on power systems*, 21(3):1458–1459, 2006.

[35] Rabih A Jabr. Adjustable robust opf with renewable energy sources. *Power Systems, IEEE Transactions on*, 28(4):4742–4751, 2013.

[36] Rabih A Jabr, Sami Karaki, and Joe Akl Korbane. Robust multi-period opf with storage and renewables. *Power Systems, IEEE Transactions on*, 30(5):2790–2799, 2015.

[37] Sunyoung Kim and Masakazu Kojima. Exact solutions of some nonconvex quadratic optimization problems via SDP and SOCP relaxations. *Computational Optimization and Applications*, 26(2):143–154, 2003.

[38] Sunyoung Kim, Masakazu Kojima, and Makoto Yamashita. Second order cone programming relaxation of a positive semidefinite constraint. *Optimization Methods and Software*, 18(5):535–541, 2003.

[39] Mikalai Kisialiou and Zhi-Quan Luo. Performance analysis of quasi-maximum-likelihood detector based on semi-definite programming. In *Acoustics, Speech, and Signal Processing, 2005. Proceedings.(ICASSP'05). IEEE International Conference on*, volume 3, pages iii–433. IEEE, 2005.

[40] Daniel Kuhn, Wolfram Wiesemann, and Angelos Georghiou. Primal and dual linear decision rules in stochastic and robust optimization. *Mathematical Programming*, 130(1):177–209, 2011.

[41] Javad Lavaei and Steven H Low. Zero duality gap in optimal power flow problem. *Power Systems, IEEE Transactions on Power Systems*, 27(1):92–107, 2012.

[42] Bernard C Lesieutre and Ian A Hiskens. Convexity of the set of feasible injections and revenue adequacy in ftr markets. *Power Systems, IEEE Transactions on*, 20(4):1790–1798, 2005.

[43] Bernard C Lesieutre, Daniel K Molzahn, Alex R Borden, and Christopher L DeMarco. Examining the limits of the application of semidefinite programming to power flow problems. In *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, pages 1492–1499. IEEE, 2011.

[44] Weixuan Lin and Eilyan Bitar. Decentralized stochastic control of distributed energy resources. *IEEE Transactions on Power Systems*, 2017.

[45] Alvaro Lorca and Xu Andy Sun. Adaptive robust optimization with dynamic uncertainty sets for multi-period economic dispatch under significant wind. *Power Systems, IEEE Transactions on*, 30(4):1702–1713, 2015.

[46] Raphael Louca and Eilyan Bitar. A hierarchy of polyhedral approximations of robust semidefinite programs. In *Decision and Control (CDC), 2016 IEEE 55th Conference on*, pages 7056–7062. IEEE, 2016.

[47] Raphael Louca, Peter Seiler, and Eilyan Bitar. A rank minimization algorithm to

enhance semidefinite relaxations of optimal power flow. In *51st Annu. Allerton Conf. Commun., Control, and Comput*, pages 2–4, 2013.

[48] Steven H Low. Convex relaxation of optimal power flow–Part I: Formulations and equivalence. *IEEE Transactions on Control of Network Systems*, 1, Issue: 1(15 - 27), 2014.

[49] Steven H Low. Convex relaxation of optimal power flow–Part II: Exactness. *IEEE Transactions on Control of Network Systems*, 1(2):177–189, 2014.

[50] David G Luenberger and Yinyu Ye. *Linear and nonlinear programming*, volume 116. Springer, 2008.

[51] Zhi-Quan Luo, Wing-Kin Ma, AM-C So, Yinyu Ye, and Shuzhong Zhang. Semidefinite relaxation of quadratic optimization problems. *Signal Processing Magazine, IEEE*, 27(3):20–34, 2010.

[52] Ramtin Madani, Somayeh Sojoudi, and Javad Lavaei. Convex relaxation for optimal power flow problem: Mesh networks. *IEEE Transactions on Power Systems*, 30(1):199–211, 2015.

[53] JR Minkel. The 2003 northeast blackout–five years later. *Scientific American*, August 13, 2008.

[54] John E Mitchell. Cutting plane methods and subgradient methods. In *Decision Technologies and Applications*, pages 34–61. INFORMS, 2009.

[55] Karthik Mohan and Maryam Fazel. Reweighted nuclear norm minimization with application to system identification. In *American Control Conference (ACC), 2010*, pages 2953–2959. IEEE, 2010.

[56] Karthik Mohan and Maryam Fazel. Iterative reweighted algorithms for matrix rank minimization. *Journal of Machine Learning Research*, 13:3253–3285, 2012.

[57] James A Momoh, SX Guo, EC Ogbuobiri, and R Adapa. The quadratic interior point method solving power system optimization problems. *IEEE Transactions on Power Systems*, 9(3):1327–1336, 1994.

[58] James A Momoh and JZ Zhu. Improved interior point method for opf problems. *IEEE Transactions on Power Systems*, 14(3):1114–1120, 1999.

[59] Daniel Munoz-Alvarez, Eilyan Bitar, Lang Tong, and Jianhui Wang. Piecewise affine dispatch policies for economic dispatch under uncertainty. In *PES General Meeting—Conference & Exposition, 2014 IEEE*, pages 1–5. IEEE, 2014.

[60] Arkadi Nemirovski, Cornelis Roos, and Tamás Terlaky. On maximization of quadratic form over intersection of ellipsoids with common center. *Mathematical Programming*, 86(3):463–473, 1999.

[61] Yurii NESTEROV. Quality of semidefinite relaxation for nonconvex quadratic optimization. Technical report, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE), 1997.

[62] James M Ortega and Werner C Rheinboldt. *Iterative solution of nonlinear equations in several variables*, volume 30. Society for Industrial and Applied Mathematics, 1987.

[63] Dzung Phan and Soumyadip Ghosh. Two-stage stochastic optimization for optimal power flow under renewable generation uncertainty. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 24(1):2, 2014.

[64] R Piwko, K Clark, L Freeman, G Jordan, and N Miller. Western wind and solar integration study: Executive summary,(wwsis) may 2010. Technical report, National Renewable Energy Laboratory (NREL), Golden, CO., 2010.

[65] Geoffrey Pritchard, Golbon Zakeri, and Andrew Philpott. A single-settlement, energy-only electric power market for unpredictable and intermittent participants. *Operations research*, 58(4-part-2):1210–1219, 2010.

[66] Benjamin Recht, Maryam Fazel, and Pablo A Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM review*, 52(3):471–501, 2010.

[67] James Renegar. Some perturbation theory for linear programming. *Mathematical Programming*, 65(1):73–91, 1994.

[68] Line Roald, Sidhant Misra, Michael Chertkov, and Göran Andersson. Optimal power flow with weighted chance constraints and general policies for generation control. *arXiv preprint arXiv:1504.00057*, 2015.

[69] Line Roald, Frauke Oldewurtel, Thilo Krause, and Goran Andersson. Analytical reformulation of security constrained optimal power flow with probabilistic constraints. In *PowerTech (POWERTECH), 2013 IEEE Grenoble*, pages 1–6. IEEE, 2013.

[70] A Santos, S Deckmann, and S Soares. A dual augmented lagrangian approach for optimal power flow. *IEEE transactions on power systems*, 3(3):1020–1025, 1988.

[71] Peter W Sauer and MA Pai. *Power system dynamics and stability*. Prentice Hall Upper Saddle River, NJ, 1998.

[72] Carsten W Scherer and Camile WJ Hol. Matrix sum-of-squares relaxations for robust semi-definite programs. *Mathematical programming*, 107(1):189–211, 2006.

[73] Alexander Shapiro. First and second order analysis of nonlinear semidefinite programs. *Mathematical Programming*, 77(1):301–320, 1997.

[74] Anthony Man-Cho So. Moment inequalities for sums of random matrices and their applications in optimization. *Mathematical programming*, 130(1):125–151, 2011.

[75] Anthony Man-Cho So and Yinyu Ye. Theory of semidefinite programming for sensor network localization. *Mathematical Programming*, 109(2-3):367–384, 2007.

[76] Somayeh Sojoudi and Javad Lavaei. Exactness of semidefinite relaxations for nonlinear optimization problems with underlying graph structure. *To appear in SIAM Journal on Optimization*, 2014.

[77] Andréa A Sousa, Geraldo L Torres, and Claudio A Canizares. Robust optimal power flow solution using trust region and interior-point methods. *IEEE Transactions on Power Systems*, 26(2):487–499, 2011.

[78] Nathan Srebro and Adi Shraibman. Rank, trace-norm and max-norm. In *Learning Theory*, pages 545–560. Springer, 2005.

[79] Tyler Summers, Joseph Warrington, Manfred Morari, and John Lygeros. Stochastic optimal power flow based on convex approximations of chance constraints. In *Power Systems Computation Conference (PSCC), 2014*, pages 1–7. IEEE, 2014.

[80] David I Sun, Bruce Ashley, Brian Brewer, Art Hughes, and William F Tinney. Optimal power flow by newton approach. *IEEE Transactions on Power Apparatus and systems*, (10):2864–2880, 1984.

[81] Joshua Adam Taylor. *Convex optimization of power systems*. Cambridge University Press, 2015.

[82] Michael J Todd. Semidefinite optimization. *Acta Numerica 2001*, 10:515–560, 2001.

[83] Reha H Tütüncü, Kim-Chuan Toh, and Michael J Todd. Solving semidefinite-quadratic-linear programs using sdpt3. *Mathematical programming*, 95(2):189–217, 2003.

[84] Luis S Vargas, Victor H Quintana, and Anthony Vannelli. A tutorial description of an interior point method and its applications to security-constrained economic dispatch. *IEEE Transactions on Power Systems*, 8(3):1315–1324, 1993.

[85] Maria Vrakopoulou, Marina Katsampani, Kostas Margellos, John Lygeros, and Goran Andersson. Probabilistic security-constrained AC optimal power flow. In *PowerTech (POWERTECH), 2013 IEEE Grenoble*, pages 1–6. IEEE, 2013.

[86] Ken McKinnon Paul Trodden W. A. Bukhsh, Andreas Grothey. Local solutions of optimal power flow problem. *IEEE Transactions on Power Systems*, 2013.

[87] Joseph Warrington, Paul Goulart, Sebastien Mariethoz, and Manfred Morari. Policy-based reserves for power systems. *Power Systems, IEEE Transactions on*, 28(4):4427–4437, 2013.

[88] Henry Wolkowics, Saigal Romesh, and Lieven Vandenrberghe. *Handbook of semidefinite programming: theory, algorithms, and applications*, volume 27. Springer, 2000.

[89] Yu-Chi Wu, Atif S Debs, and Roy E Marsten. A direct nonlinear predictor-corrector primal-dual interior point algorithm for optimal power flows. *IEEE Transactions on power systems*, 9(2):876–883, 1994.

[90] Liang Xie and Hsiao-Dong Chiang. A enhanced multiple predictor-corrector interior point method for optimal power flow. In *Power and Energy Society General Meeting, 2010 IEEE*, pages 1–8. IEEE, 2010.

[91] Liang Xie and Hsiao-Dong Chiang. Weighted multiple predictor-corrector interior point method for optimal power flow. *Electric Power Components and Systems*, 39(2):99–112, 2011.

[92] Yinyu Ye. Approximating quadratic programming with bound constraints. *Mathematical Programming*, 84:219–226, 1997.

[93] Yinyu Ye. Approximating quadratic programming with bound and quadratic constraints. *Mathematical Programming*, 84(2):219–226, 1999.

[94] Baosen Zhang and David Tse. Geometry of feasible injection region of power networks. In *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, pages 1508–1515. IEEE, 2011.

[95] Shuzhong Zhang. Quadratic maximization and semidefinite relaxation. *Mathematical Programming*, 87(3):453–465, 2000.

[96] Ray D Zimmerman and Carlos E Murillo-Sánchez. Matpower 4.1 users manual. *Power Systems Engineering Research Center (PSERC)*, 2011.